VOL. 3, NO. 1, MARCH, 1961

human factors

ERSITY
RY

THE JOURNAL OF THE HUMAN FACTORS SOCIETY

OHIO STATE U
EDUCATION

Contents

SOME IMMEDIATE EFFECTS OF DROWSINESS AND SLEEP ON NORMAL HUMAN PERFORMANCE 1
Charles W. Simon

A TECHNIQUE FOR EVALUATING MAN-MACHINE SYSTEM DESIGNS 18
Arthur I Siegel and J. Jay Wolf

OPERATOR ACCURACY IN ANGLE MEASUREMENT AND TRANSFER 29
Robert C. Haygood and Richard K. Overton

ANTHROPOMETRIC CONSIDERATIONS FOR ESCAPE CAPSULE DESIGN 36
John W. Chaffee

AN EXPERIMENTAL INVESTIGATION OF THE INTERACTION BETWEEN PROBLEM LOAD AND LEVEL OF TRAINING 53
Richard A. Behan, C. Russell Bughman, James N. Bumpus, and Sally Bonney Gilbert

COED—A DEVICE FOR THE EXPERIMENTAL STUDY OF MAN-MACHINE SYSTEMS 60
Robert H. Johnson, Donald A. Gordon, Bruce O. Bergum, and Wayne E. Patterson

OPERATIONAL SEQUENCE DIAGRAMS IN SYSTEM DESIGN 66
Martin I. Kurke

THE AUTHORS 74

THE MACMILLAN PRESS NEW YORK • LONDON • PARIS • LOS ANGELES

T58
A2 H8
V 3
1961

OFFICERS AND EXECUTIVE COUNCIL

The Human Factors Society is a non-profit California Corporation. The officers and members of the Executive Council are:

PRESIDENT: Stanley N. Roscoe, 1961
PRESIDENT-ELECT: Jack W. Dunlap, 1961
VICE-PRESIDENT: Wesley E. Woodson, 1961
PAST-PRESIDENT: Arnold M. Small, 1961
SECRETARY: Paul G. Cheatham, 1961
TREASURER: Donald W. Conover, 1961
EDITOR-IN-CHIEF: Stanley Lippert, 1959–1961

EXECUTIVE COUNCIL: Charles I. Barron, 1960–1962; Max W. Lund, 1959–1961; John Lyman, 1960–1962; Ross A. McFarland, 1959–1961; John J. Swearingen, 1961–1963; Lawrence Slote, 1961–1963

EDITORIAL BOARD: Walter P. Grether, Blaine H. Levedahl, J. C. R. Licklider, Stanley Lippert, John Lyman, John P. Meehan, Allison K. Simons, Ezra V. Saul, Neil D. Warren, Joseph W. Wulfeck

HISTORIAN: Donald T. Hanifan

Address general correspondence to Human Factors Society, 4113 Lee Highway, Arlington 7, Virginia, Attention: **Paul G. Cheatham**, Secretary, for Society business including payment of dues.

Address matters pertaining to the Journal or Bulletin to **Human Factors**, Biotechnology Laboratory, Department of Engineering, University of California, Los Angeles 24, California, Attention:

John Lyman, Managing Editor, for Journal or Bulletin Subscriptions or other publication business, including changes of addresses.

Stanley Lippert, Editor-in-Chief, for submission of Journal articles and letters to the editor.

Hilde Groth, Bulletin Editor, for Bulletin news.

COPYRIGHT © 1960
by the HUMAN FACTORS SOCIETY

PUBLISHED BY
PERGAMON PRESS, 122 E. 55th St., New York 22, N.Y.
Headington Hill Hall, Oxford, England.

Printed in Great Britain.

SUBSCRIPTION PRICE: MEMBERSHIP in Society includes subscription to both publications.

Journal: NON-MEMBERS SUBSCRIPTIONS TO BE SENT DIRECT TO THE PUBLISHER, PERGAMON PRESS. Commencing with Volume 2 the subscription price will be:
$10.00 (£3 10s.) per annum, for individuals who certify that the journal is for their personal use. $20.00 (£7) per annum, for libraries, government establishments and research institutions.

Bulletin: $5.00 per year. Orders to be sent to John Lyman.

Some Immediate Effects of Drowsiness and Sleep on Normal Human Performance,[1]

CHARLES W. SIMON, Hughes Research Laboratories, Malibu, California.

Man spends the greater portion of his life in sub-optimum states of consciousness. The paper reviews the experimental literature describing normal human performance at different levels of drowsiness and sleep, as these states are defined by electroencephalographic patterns. Among the types of behavioral responses to stimuli discussed are reflexes, simple movements, detection (i.e. arousal to external stimuli and awareness of internal states), discrimination, learning, and recall. Non-behavioral electroencephalographic responses are also described. Suggestions are made for future research.

INTRODUCTION

One-third of a man's life is spent sleeping. An even greater proportion of the remainder of his life is spent in below optimum states of consciousness. To be able to understand, predict, and control the effects that these 'subacon' states have on human behavior is of importance to nearly everyone for effective day-to-day living. Many practical tasks encountered in business, education, medicine, and the military are directly affected by this relationship between levels of drowsiness and sleep and human performance.

A few general questions will help to illustrate the breadth of this problem area:

What effect can we expect drowsiness (from any cause) to have on the quality of decisions a man in a key position makes? Will the tired business man or the exhausted military commander be less critical, show poorer judgment, and fail to make optimum use of the existing knowledge because of his neurophysiological state?

How can we design a warning system which not only arouses an individual when the situation becomes dangerous, but which also alerts him when he is no longer capable of performing effectively?

What minimum level of alertness is necessary for particular jobs when it is known that, because of the nature of the situation, alertness cannot be optimum? How can the equipment be designed to optimize operator alertness to meet the task?

Can learning during the drowsy period prove to be a practical adjunct for improving waking learning? Can this be used as a means of providing additional instruction to the child at school, or the soldier, or the industrial employee being trained for a technical job?

Can a man be permitted to sleep on a job between critical periods and be alerted by significant external cues to which he has been conditioned to act?

In this paper, a survey will be made of the relatively scant literature which relates neurophysiological states along the wake-sleep continuum to behavior. Relatively few papers will be included which infer the levels of drowsiness or sleep of the subjects being studied. Instead, the papers to be included will generally define these levels operationally by characteristic *electroencephalographic patterns* associated with them. Examples of these are shown in Fig. 1. For the moment, it shall be stated categorically that the EEG appears to be the best single criterion of drowsiness and sleep levels that exists. This conclusion will be discussed in more detail in Appendix A.

Many studies of sleep and the EEG have been done with lower animals. The reason for this is obvious. Most of the people doing work in this field have been concerned with the

[1] This paper was presented in December, 1959, at a Symposium on Sleep Research, a program of the Department of Psychology, University of Florida, sponsored by the U.S. Air Force Office of Scientific Research.

physiology, the structure, and the mechanisms underlying sleep and the EEG. Cutting and probing the human brain is dangerous and can be done only under unusual circumstances; therefore, lower animals have been used. Since a great deal of the behavior discussed here deals with complex mental processes, generalizations from the animal work are unsafe. So, except when it appears appropriate to do otherwise, the studies described shall be confined to experiments with *human subjects*.

Much of the work on human subjects has been clinical in nature, emphasizing the abnormal rather than the normal man. Interest in the EEG during sleep has grown out of its potential applications for the detection and diagnosis of brain disorders. However, the papers surveyed here will deal primarily with the *normal* human adult.

Finally, many of the studies on sleep have used drugs to create sleep artificially. However, it is important to understand the effects of *natural sleep* in man; studies which include artificially induced sleep will be ignored at this time.

Thus, the papers to be reviewed will be concerned with the performance capabilities and neuro-physiological correlates of normal humans during natural conditions of drowsiness and sleep.

It is of interest to note that a large portion of the research on sleep has been done outside of the United States. A count of the papers listed in *Current List of Medical Literature*[6] under the category "sleep" from 1955 to 1959 inclusive yielded the following facts:

1. During this five-year period, approximately 550 titles were listed.
2. In 1955, 188 titles were listed; in 1959, only 68. There was a steady decrease over this period.
3. Of the 550, only 21 per cent were published in the United States. Forty-six per cent were published in Russia or in their satellite countries, primarily Poland and Czechoslovakia. Fourteen per cent were published in Germany and 13 per cent in France.

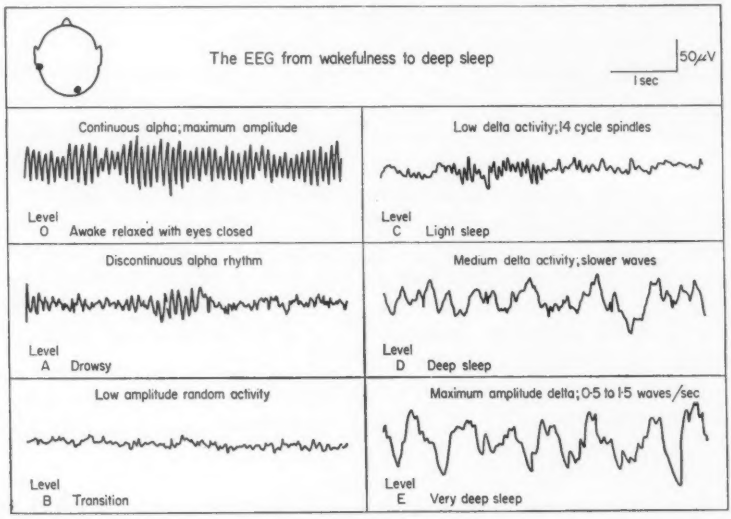

The EEG from wakefulness to deep sleep
50μV
1 sec
Continuous alpha; maximum amplitude
Level O Awake relaxed with eyes closed
Low delta activity; 14 cycle spindles
Level C Light sleep
Discontinuous alpha rhythm
Level A Drowsy
Medium delta activity; slower waves
Level D Deep sleep
Low amplitude random activity
Level B Transition
Maximum amplitude delta; 0·5 to 1·5 waves/sec
Level E Very deep sleep

Fig. 1. The EEG from wakefulness to deep sleep. (Adapted from Fig. 1 of Ref. 13, courtesy of the *American Journal of Psychology*.)

4. Approximately 40 per cent of the articles pertained to therapeutic uses of sleep, and around 35 per cent were concerned with physiology.

Relatively few foreign studies have been included in this paper since a majority of them fail to meet the criteria established above.

SIMPLE SENSORY-MOTOR BEHAVIOR

It is obvious that all bodily activities do not cease during sleep. The vegetative, or autonomic functions, necessary for supporting life continue, although generally at a reduced rate. Thus during sleep, heart rate and respiration rate are reduced, body temperature drops slightly, gastric contractions are essentially unchanged, while cranial blood volume increases. However, it is the behavior associated with the central nervous system in which we are interested in this paper.

In this section, we will concern ourselves with the performance of simple sensory and motor tasks at different levels of drowsiness and sleep. In the first part, unconscious motor behavior is emphasized, while in the second part the tasks involving simple awareness of stimuli are discussed. In the third part, neurological responses to stimuli are described.

Musculature Changes and Simple Motor Responses at Different Levels of Drowsiness and Sleep

Let us examine some of the changes in muscle tone and simple, involuntary motor responses that have been associated with different levels of drowsiness and sleep.

Muscle tone. Perhaps one of the most general observations relating body tone to drowsiness and sleep is that as one grows sleepy one tends to relax, and conversely, to grow sleepy, one needs to relax. Blake and Gerard[4] found that when a subject is told to hold a spool or a piece of paper between his fingers, he can do so as long as alpha activity is observed in his EEG. When the subject becomes drowsy and the alpha becomes intermittent and disappears for even a few seconds, his muscle tone diminishes and he drops the objects. This technique has been used as a criterion for the onset of sleep when no EEG's can be taken.

Reflexes. Reflexes are among the simplest of motor responses to stimuli. Generally, no cortical involvement is required to achieve reflexive responses to unconditioned stimuli. Morgan and Stellar[26] write: "In general, it is harder to elicit reflexes during sleep than in the waking state. This characteristic does not apply equally well to all reflexes, however. While the threshold for eliciting proprioceptive reflexes is heightened, the cutaneous reflexes are nearly normal. The knee jerk, for example, gradually disappears as a man goes to sleep. Cutaneous stimulation, on the other hand, seems to bring out rather normal responses. Brushing the face produces grimaces or tickling stimuli may elicit scratching movements. Of particular interest is the fact that the Babinski reflex may make its appearance in sleep. This reflex, the fanning of the toes, can be elicited in infants by scratching the sole of the foot. It never appears in the waking adult, however, unless there is a neural injury that interferes with the flow of impulses through the pyramidal pathways from the precentral gyrus of the cortex. The pupils tend to get smaller as sleep deepens, but they still respond to light until the very deepest sleep is reached. The threshold for pain also diminishes during sleep."

Body movements. Gross body movements which occur during drowsiness and sleep may occur in response to external signals or from internal somesthetic stimulation such as the constrictions in circulation brought about by lying in restricted positions.

The frequency with which body movements occur has been used as one measure of the depth of sleep. The fewer the movements, the deeper the sleep is assumed to be. Some observers believed that these movements can be initiated at practically every level of sleep the subject is in, but less frequently at the lower levels. However, Hoffman *et al.*[17] report

that body movements were regularly associated with a change in the EEG to a pattern suggesting lighter sleep. In most cases the alteration actually preceded the movement. Dement and Kleitman[10, 11] have observed a similar shift upward from all levels of sleep, particularly with major movements.

The muscle potentials accompanying movements often obscure the brain potential records. Simon and Emmons[31], using "eyeball" scoring techniques, found that when the alpha frequencies were observed during or at the end of a record hidden by movement artifacts, factual material presented at that time tended to be heard and remembered more often than when no alpha was observed. They concluded that the presence of alpha and not movement was the critical criterion for conscious responses.

The nature of eye movements seem to vary with depth of sleep. Aserinsky and Kleitman[2] were able to relate the presence of dreams with eye movements in the borderline states. These eye movements were rapid, jerky, and binocularly symmetrical. In deep sleep, however, eye movements are slow and rolling and appear to be less under the control of the subject and do not appear to be associated with dreaming.

Summary. General body tonicity and muscular activities are reduced in proportion to the depth of drowsiness and sleep. Reflexes drop out differentially as one goes to sleep. Most of the studies of these relationships have been piece-meal and descriptive rather than broad in coverage and quantitative. A systematic study of reflexive responses classified on the basis of nervous system and skeletal-musculature involvement along with precise measurements of the physical characteristics of stimuli presented at all levels of drowsiness and sleep (as measured by the EEG) would prove to be a valuable area of investigation.

Arousal to External Signals and Awareness of Internal States During Drowsiness and Sleep

It is a common observation that as one becomes drowsy, the more unlikely will be a detection of and voluntary response to external signals. With drowsiness, we are less alert and do not attend as diligently to the particular designated signal. If we do detect it, we are often slower in responding. Furthermore, once asleep, the more soundly we sleep, the more difficult it is for us to be aroused. It is on the basis of this last assumption that the EEG patterns have been related to depth of sleep by numerous experimenters. Unfortunately, although many of the early studies ordered the EEG patterns along a depth of sleep scale, quantitative data about the stimuli were seldom supplied.

Arousing to tones. Loomis, Harvey and Hobart[23, 24] were among the first to report in detail the relation between EEG patterns and depth of sleep. Based on the relative intensity of tones and light necessary to arouse sleeping subjects enough to squeeze a bulb, they were able to identify a variety of EEG patterns at different levels of wakefulness, drowsiness, and sleep. In general, these EEG patterns were similar to those shown earlier in Fig. 1.

Blake and Gerard,[4] when a particular EEG pattern was clearly present, would sound a 1000-cycle tone of constant intensity, and then softly ask the subject, "Are you awake?" at 10-sec intervals for a maximum of one minute or until the subject answered verbally. The more delayed the awakening, the deeper they assumed the sleep to be. They found the same general relations between the EEG and levels of drowsiness and sleep as Loomis *et al.* had.

Blake, Gerard and Kleitman[5] correlated with EEG potentials the duration of a constant intensity tone required to elicit the bulb squeezing response from their subjects during a 6-hr sleep period. Their results were in essential agreement with Loomis *et al.* They described subjects as being awake when alpha plus delta plus 14 c/s frequencies were present and in deep sleep when delta frequencies (and occasionally 14 cycles) were present. During the latter part of the sleep period, subjects were in a light sleep when null or low voltage waves were present. Upon awakening, intermittent

alpha was again observed. Although not actually specified, it was implied that the duration of sound from onset to arousal corresponded to the sleep levels. They stated that "loss of consciousness and of alpha rhythm are related".

Self awareness. The relation between EEG patterns and the subject's judgment of his own state of wakefulness, drowsiness or sleep has been investigated. The subject's introspections are consistent with the other data relating detection capabilities to EEG patterns. Thus, Davis, Davis *et al.*[7] requested the subject to signal if he had "dozed off". They noted that alpha would disappear and then return for a period of 3 to 23 sec before the subject would signal that he had dozed and returned to consciousness.

Blake, Gerard and Kleitman[4] would quickly arouse their subject and ask him if he had been asleep. Again the relation between alpha and the subject's subjective impression of his state of consciousness followed the same pattern. When he would report he had been awake, the prior 5-min period would have had continuous alpha and no delta. If he reported dozing with some dimming of awareness and a decreased sense of reality of surrounding events, alpha waves would have been absent for at least 3 and an average of 6 sec, with no delta. When light sleep was reported, the average period without alpha was 9 sec, and when deep sleep was reported, the average period that had been without alpha was 55 sec. Thus the longer the period without the alpha frequencies, the more drowsy or asleep the subject said he had been.

Summary. It is well established that there is a high and consistent relationship between the introspective description of one's state of drowsiness and sleep and the frequency and amplitude of the EEG's. Similarly, the intensity and duration of stimuli required to arouse an individual from some level along the wake-sleep scale correspond with the relation of alpha and delta cortical potentials. What is needed is a systematic study to relate the physical characteristics of the stimuli for different sense modalities with the particular EEG patterns representative of different levels of drowsiness and sleep and with quantitative measures of responses of varied complexities.

The Effect of Stimulation on the EEG During Drowsiness and Sleep

The individual who is conscious and otherwise physically capable should be able to indicate verbally or by overt movement that he has detected a signal to which he has been told to attend. Because he fails to make an overt response does not necessarily mean that the stimuli were not or could not be detected. Loomis, Harvey and Hobart[23, 24] reported that when alpha had disappeared almost completely and no delta was present, subjects could be observed unsuccessfully trying to make the response of squeezing a bulb. Reduced muscle tone at this borderline of sleep could account for this. Simon and Emmons[29] and others have reported that when a subject became quite sleepy, he would purposely fail to respond as instructed in order not to disturb his rest.

Before discussing further the behavioral responses which can occur at various levels of wakefulness and sleep, it is important to mention changes in the brain waves that appear as *effects* or responses to stimulation. Certain momentary changes of the EEG patterns—in addition to those used as a criterion of depth—occur at various levels of wakefulness and sleep as involuntary responses and indicate that impulses created by the stimulus have reached the brain. The characteristics of these cortical responses differ at different levels of wakefulness and sleep.

Alpha blocks. Among the more common of these response phenomena at the waking end of the continuum is the "alpha block". Strong continuous alpha will reduce to low amplitude random activity when the eyes of most subjects are opened, and will reappear when they are closed. This is not the intermittent alpha mentioned earlier, but an effect definitely related to stimulation. Although visual stimuli (e.g. light) have always appeared the most effective for depressing alpha rhythms, blocks are readily

observed to occur to sounds as well as electric shock and other tactile stimuli. After continuous stimulus presentations for long periods of time, the alpha block phenomenon has been found to adapt and not to occur upon stimulation.

It is generally agreed today that the occurrence of alpha blocks is closely related to the attention or arousal value of the stimuli. Jasper and Cruikshank[18] stated that stimuli depressed the alpha by virtue of their signal value. When a subject is attending a stimulus, the probability of a block increases. Milerian[25] reported that alpha blocks have been observed when the stimulus attracts the attention of the subject, when the stimuli are to be memorized, when they occur during emotional states, and when solutions to mental problems are being sought. Loud sounds may not produce an alpha block if the subject does not attend, while a low sound requiring concentrated attention may produce one.

The alpha block is observed to occur in response to stimulation of the fully awake and relaxed subject. When drowsiness occurs, however, and intermittent alpha and considerable low voltage activity can be observed, the alpha no longer blocks upon stimulation but tends to return in bursts.

Conditioning alpha rhythm. As in the case of motor responses to stimulation, alpha block responses that indicate the detection of a signal are subject to conditioning. Considerable work of this type has been done both in the United States and in Russia.

Jasper and Shagrass[19] were able to obtain a conditioned alpha block to a sound which had previously had no appreciable effect. Approximately ten trials were required for conditioning which was extinguished (without reinforcement) in three to five trials. They also were able to establish delayed, trace, differential, and backward conditioning.

Of particular interest to our discussion of sleep were the delayed and trace conditioning. In the former, the sound was presented continuously for 10 sec overlapping the appearance of the unconditioned stimulus, the light. For trace conditioning, the sound was turned off but the light was not presented for approximately 10 sec. In both types of conditioning, alpha blocks occurred to the sound stimulus after delays of from 9 to 45 sec. Extinction generally occurred within four trials without reinforcement. The interesting thing is that occasionally with delayed and trace conditioning, alpha would disappear (not block) and slow delta waves would appear during the unreinforced delays, and two subjects who failed to give the usual manual response indicated they had dozed off.

Pavlov observed this in dogs and also in humans when shock was used to condition a manual response to a tone.

Subliminal detection. Rusinov[28] reports that a number of authors disclosed the possibility of developing conditioned reflexes to acoustic stimuli, the intensity of which was lower than the threshold for verbal acknowledgment.

The establishment of conditioned reflexes to subthreshold stimuli requires a considerably larger number of training combinations than that to superthreshold stimuli; the latency of the conditioned alpha rhythm is also much longer for subthreshold stimuli.

This work has been done using alert subjects. It should be noted that although this form of subliminal detection and detection during drowsiness and sleep both involved responses to stimulation lower than the threshold for verbal acknowledgment, subliminal conditioning takes place when the stimulus is weak but the sensors are normal, while for conditioning during sleep, the stimuli are normal but the sensors or some connecting mechanisms are operating below par.

Evoked potentials: on-effects and K-complexes. The only true electroencephalographic responses to stimulation which are readily observable at all stages from wakefulness through deep sleep are those associated with the "on-effects". P. A. Davis[9] has described these potentials evoked by stimulations. During the drowsy states, the on-effect, composed

of diphasic and sometimes triphasic waves is most prominent in the vertex. The first phase is negative. The total duration of the on-effect is approximately 0.3 sec or less. The voltage, measured from the peak to trough, ranges from 100 micro-volts to deflections which are just distinguishable from the background of physiological activity. The latencies are of the order of 30 to 40 msec. On-effects are more readily seen in the records of non-alpha dominant subjects because they tend to be hidden in the higher amplitude patterns of alpha-dominant subjects.

On-effect responses to sound stimuli take place in a variable manner. They do not appear to be related to loudness or frequency in the range of the tones used. As with alpha blocks, on-effects were related more to the value of the stimulus to the subject, occurring more often when the subject was uncertain or when he anticipated the stimuli.

Hernandez-Peon and Donoso,[15] using five patients, found that cortical potentials evoked by flashes of light at a constant intensity could be recorded by deeply implanted electrodes in the occipital lobe. The magnitude of the evoked potential varied with the degree of attention to the lights. If the subject attended to arithmetic problems rather than the lights, evoked potentials were reduced in amplitude or disappeared. Jouvet and Courjon[20] found a similar effect. They had subjects attend by counting flashes but when they attended to another sense modality there was a diminution in the amplitude of the evoked visual response.

P. Davis[9] wrote: "If the physiological condition of the subject becomes modified as he shifts from the alert to sleep, and the 'psychological set' is maintained, the on-effect becomes more pronounced and more predictable." During sleep, the on-effect is evident with a return of fast waves; this is the onset of the K-complex.

H. Davis *et al.*[8] have described the K-complex in some detail. It has been found to occur after a tone stimulus, first as an electrically negative potential of some 50–100 mV, which after about three-quarters of a second, becomes positive by 100 mV or more. This is often followed by original electrical levels, or fast waves of the 14 cps variety are found superimposed on the slow waves. K-complexes may appear once or in a sequence, and they adapt with repetition. They appear in the sleep states and are not associated with arousal. In fact, when subject is aroused, the K-complex is no longer observed. They appear primarily in the vertex, although they spread over the entire head. They appear as non-specific responses, light and other stimuli will all cause them to appear, and they sometimes appear spontaneously in response to the snoring of the subject. Attempts to condition K-complexes have not been too satisfactory. Their presence shows that signals are reaching the cortex even in the deeper sleep levels; however, the quantity of the stimulus plays the important role, not the quality. K-complexes do not indicate that the quality of the stimulus has been maintained, or can be distinguished.

Summary. Since it is possible to evoke responses in the EEG when other overt responses are lacking, the different circumstances concomitant with these must be studied. The apparent relation between the attention value of the stimuli and these cortico-electrical responses bears on the problem of discrimination to be discussed more in the next section. EEG-evoked potentials do not appear to seriously affect the use of the EEG as a measure of depth of sleep; however, more work is still needed on this problem in order to refine and understand existing measures.

HIGHER MENTAL PROCESSES

Although the ability to make voluntary motor responses and detect sensory stimuli are critical aspects of consciousness, the level need not be high. It is the higher mental processes that are ultimately of interest to us—the ability to make decisions and to solve problems. This type of performance comes about through the

complex integration and differentiation of the incoming stimuli with our past experience and involves an interplay of many aspects of the nervous system. Only when the level of consciousness is high is performance optimum.

Unfortunately, most of the work investigating these higher mental processes has been more clinical than experimental, more descriptive than analytical. They have often been a result of investigations of the behaviors of the psychotic or of ablated surgical specimens. In normal man, except in the studies of the effects of sleep deprivation[34] few complex behaviors have been related with the cortical potentials of drowsiness and sleep.

In the next two sections, the studies to be described will relate levels of drowsiness and sleep (and the EEG) with (1) the ability to make immediate discriminations among stimuli and (2) with the ability to learn or remember.

Discriminations at Different Levels of Drowsiness and Sleep

There appear to be many situations for which can be postulated an arousal system which can be selectively activated by previously conditioned responses to complex stimuli. The most popular anecdote of selective discriminations being made during sleep has been that of the arousal of the mother to the cry of her baby but not to the sound of the train. This phenomenon has never been scientifically verified. Li, Jasper and Henderson,[22] while studying various pathological forms of sleep, described one narcoleptic patient who would sleep through a loud noise produced by striking a brass bucket, but would awaken when her name was called. Interestingly enough, the EEG responded to both the sound of her name and the bucket. That the differences observed in cortical potentials depend on the "attention getting" value of the stimulus has already been described. However, the characteristics of the stimulus that make it "attention getting" have not been adequately indicated, and surprisingly enough, little definitive experimental work has been done on this problem.

Discrimination during sleep. Rowland's[27] study with cats represents one of the few systematic investigations of this problem. Cats were conditioned while they were awake to the sound of a tone, with shock being used as the unconditioned stimulus. The cats' responses to the shock and the conditioned tone was to "flinch" and otherwise show fear responses. Rowland also introduced some neutral tones during the conditioning process to ensure obtaining differential responses. After keeping the cats awake the day before, they were then permitted to sleep.

During sleep the now conditioned tone was presented again, but never with the shock. The following results were observed:

1. During sleep, flinching was not observed to the tone as it had been during the waking state.
2. However, when the conditioned tone was presented, the EEG activity showed a change from patterns representing sleep to the low voltage fast activity associated with arousal.
3. The first time that a neutral tone was presented during sleep, the EEG showed a similar change; however, there was a rapid adaptation to the neutral tone but not to the conditioned tone.
4. Although both motor and cortical responses to the conditioned tone were run to extinction during sleep, the flinch response was observed to reappear in the waking state.

In a series of tests, Rowland showed that the cats could distinguish during sleep (as reflected by their EEG's) sounds of different quality, tones of different frequencies, different number of stimuli, and the like.

Behavioral discriminations during drowsy states. As previously stated, most of the work examining the complex mental functions during drowsy states has been done by those investigating the effects of sleep deprivation. Since the work in that area is quite extensive, very little will be described here except to illustrate how it fits

into the general picture of consciousness, sleep, and the EEG. The only real distinction between sleep deprivation studies and those presented in this paper is that in the former the drowsiness state is caused by being deprived of normal sleep for abnormally long periods. The exact effect this has on our measures is not completely certain. Also, whereas we have been concerned with the effects of levels of drowsiness and sleep on performance, sleep deprivation brings in the dimension of time and the distribution of periods of work and rest.

The brain wave patterns of the sleep deprived person are those of a drowsy individual who may doze off occasionally.[3, 34] This comes as no surprise. The more recent investigators, observing the intermittency of the alpha rhythm, have now associated it with drops in performance attributed to "lapses". Armington and Mitnick[1] describe just such a relationship. Their sleep deprived subjects showed a decline in the amplitude of the alpha waves relative to the number of hours of sleep lost. When subjects were required to distinguish a letter "X" from others and to indicate this by moving a lever, errors of omission consistently exhibited a lower amount of alpha activity concurrent with the response than did correct responses or errors of commission.

Bjerner,[3] also using sleep deprived subjects, noted similar relationships among the cortical potentials and performance on a discrimination task. His subjects were required to move a lever in one direction when they heard a buzzer. The less alpha activity that was present, the slower the response. When slow delta waves were observed, subjects sometimes failed to respond, or did so only after extreme delays.

Cortical potential discriminations. In the sleep loss studies, the brain waves have generally been used as a criterion associated with drowsiness and sleep and not as a response to the stimulation. With Rowland's animal study, it was observed that discriminations continued to be made by the EEG, even in sleep, after behavioral discriminations no longer occurred. In the sleep loss studies described, there was evidence of a slowing or absence of overt responses, but no investigation was made to see if discriminations occur cortically. More work with human subjects is required on the problem.

Durup and Fessard[12] have noted that the onset of an alpha block is faster when the subject is required to discriminate between two tones than to detect one, although the motor response was slower. This fits the general pattern of the alpha response being related to attention and mental effort, and does not contradict the other work described.

Summary. When behavioral responses are used to discriminate among stimuli, the responses occur less frequently and more slowly as drowsiness becomes deeper and they disappear completely during sleep. Drowsiness is associated with a decrease in proportion, amplitude, and frequency of the alpha rhythm and sleep is noted by the appearance of slow non-alpha (i.e. delta) waves.

When the observed response is a change in the EEG, however, it is reported that discriminations continue to be made between stimuli previously conditioned to be important to the subject. That these are conscious discriminations is unlikely and the practical implications are still unknown.

A systematic investigation of the discriminative capacity of the normal human in natural states along the drowsy-sleep continuum could be a significant contribution toward understanding the psychology and the neurophysiology of consciousness and sleep. But it means investigating every sense modality. It requires dimensionalizing stimuli—their physical composition, complexity, content, and meaning to the individual. It further requires the dimensionalizing of responses in terms of their neurophysiological complexity; and finally, it demands that cortical potentials and other critical electrochemical measures be taken. Association of the muscular effort and metabolic costs for the performance of different tasks at the different wake-sleep levels is an important adjunct in these studies of the more complex mental processes.

Learning and Remembering Material Presented at Different Levels of Drowsiness and Sleep

Memory and learning are both complex concepts. Most persons accept the hypothesis that the ability to remember and to learn is indicative of an aspect of conscious behavior. In the previous section we discussed the ability to discriminate, while drowsy or asleep, conditions which had been established previously in waking states. Now we will examine the effects of being drowsy or asleep on the ability to remember and learn. Four types of conditions will be considered: the ability to remember events that are internally generated, the ability to recall external stimuli immediately after the occurrence, the ability to develop simple conditioning during sleep, and the ability to learn complex verbal material during drowsiness and sleep.

Memory of dreams. Dreams are presumed to occur in the borderline states between wakefulness and sleep where alpha frequencies have slowed and the amplitude and proportion of alpha rhythms are essentially zero (Dement and Kleitman[10, 11]). At these levels, the awareness of external stimuli has diminished and dreams, the lowest level of consciousness, are dominant.

Blake, Gerard, and Kleitman[5] awakened their subjects suddenly and asked them if they had been dreaming. A direct relation was observed between the average time alpha was absent and the ability to recall the dreams. When alpha had been absent for 9 sec and no delta was present (drowsy state), dreams and their contents were remembered. When alpha waves had been absent for an average of 55 sec with some delta activity present (borderline state), subjects recalled that dreams had occurred but did not recall their content. When prominent delta waves were observed (sleep state), subjects could not recall having dreamt at all.

Immediate memory. Hoagland[16] studied both normal and also cerebral arteriosclerotic patients (mean age, 75 years), relating the median frequency of their waking EEG with their ability to recall immediately a series of digits in reverse, to repeat immediately a simple sentence, and to recall immediately short stories. He found a significant correlation of 0.88 between this immediate memory and the median EEG frequency. The slower the median frequency, the poorer the memory for recent events. This is of interest in our study of the relation between memory and sleep because the slowing of alpha frequencies has been shown to be characteristic of an increase in drowsiness. Hoagland found no correlation between these median frequencies and the subject's memory for remote events.

Conditioning during sleep. Krasnogorski[21] attempted to form a salivary conditioned response in a child who was asleep. He introduced as the unconditioned stimulus a concentrate of citric acid into the child's mouth through a tube. This attempt to condition was unsuccessful.

In this case, no electroencephalographic records were made. It has been pointed out that behavioral responses do not show up readily during sleep; choosing another reflex to condition may have made a difference.

Memory for factual material. Simon and Emmons[29, 30, 31] studied the ability of normal adult males to retain the answers to factual questions presented at different levels along the wake-sleep continuum. Twenty-one subjects were chosen with average or above IQ's and with a monopolar occipital EEG showing a continuous alpha rhythm when they were awake but resting with their eyes closed. After being pretested to see whether or not they knew the answer to 96 factual questions on history, sports, science, and the like, they retired to spend a normal 8-hr night's sleep. The same 96 questions along with the correct answers were played once each at 5-min intervals during the night. Continuously throughout this entire period, monopolar EEG recordings were made from each subject's right occiput and vertex. A pen marker showed the exact sections of the EEG record during which the questions and the answers occurred. After the 8-hr sleep

training period, subjects were awakened and given only the questions again and tested to determine which of the answers not known previously could now be recalled unaided. With this data, it was possible to classify the answers into groups according to the EEG pattern present during the answer periods and to determine the percentage of correct answers associated with each pattern.

Since alpha-dominant subjects were used, it was possible to classify by eye EEG patterns in the A states to a finer degree than had been done in many of the previous classifications. These A patterns are shown in Fig. 2. The other patterns in sleep states are the same as those shown in Fig. 1.

The results of each individual subject were consistent with that of the group. Within waking level 0, approximately 8 per cent fewer items were recalled unaided than had been recalled by another group that had not only been awake but sitting up and trying to remember. At this level, a slight reduction in the amplitude of the continuous waking alpha before going to sleep and after awakening from sleep was related to a similar decrease in the probability that an appropriate answer would be given. As the percentage of alpha decreased in quantity and amplitude in the A+ and A levels, there was a corresponding decrease in the probability of recall. As the individuals became very drowsy and the slower frequency (A−) patterns were observed, still fewer items were recalled. When the alpha patterns were absent (within 30 sec on either side of the answer period (B)) and delta activity (C, D, and E, Fig. 1) was dominant, no recall was observed.

Following the unaided recall test, subjects were given the questions again along with five possible answers to see if items could be recognized during sleep even if they could not be recalled unaided. However, the same general trend previously observed occurred with some qualifications. As would be expected, the percentage of answers recognized correctly was always higher than those correctly recalled unaided. Furthermore, on the recognition test, approximately 23 per cent of those answers supplied when delta activity was observed in the EEG's were recognized. This did not mean that memory as measured by a recognition test for materials presented during sleep was possible, for it was found that a control group, not supplied with any answers whatsoever, also correctly recognized 23 per cent of the answers on the multiple choice portion of the test, which of course is approximately what would be expected to have occurred by chance on a five-item selection.

In a second study by Emmons and Simon[13], ten one-syllable nouns were repeated as many times as possible to nine subjects during an 8-hr sleeping period. The mean number of repetitions for all words was forty-six times. A continuous EEG recording during the presentation of the training material was used to determine the sleep level at that time and the material was turned off as soon as movement or cyclical activity within the alpha frequencies was observed. Seventy-four per cent of the repetitions occurred in borderline or light stages of sleep. Experimental subjects did not do significantly better than the controls in selecting (from a list of fifty words) the training words presented during sleep. Nor did they select the training words any more frequently than they did a randomly selected list of ten words.

EEG's affected by the stimulus. In the first study reported by Simon and Emmons,[31] approximately 14 per cent of the items were associated with changes in EEG patterns in response to the auditory stimulation (Fig. 3). Thus, the continuous alpha was blocked at times during the period when the answer was being given. When the EEG was affected by answers given during the null states (B), (where neither alpha nor delta activity was observed, i.e. the transition level between wakefulness and sleep) very high frequency, low amplitude activity (which has been associated by other researchers with arousal) was observed. And in the sleep states, the stimulus activated the

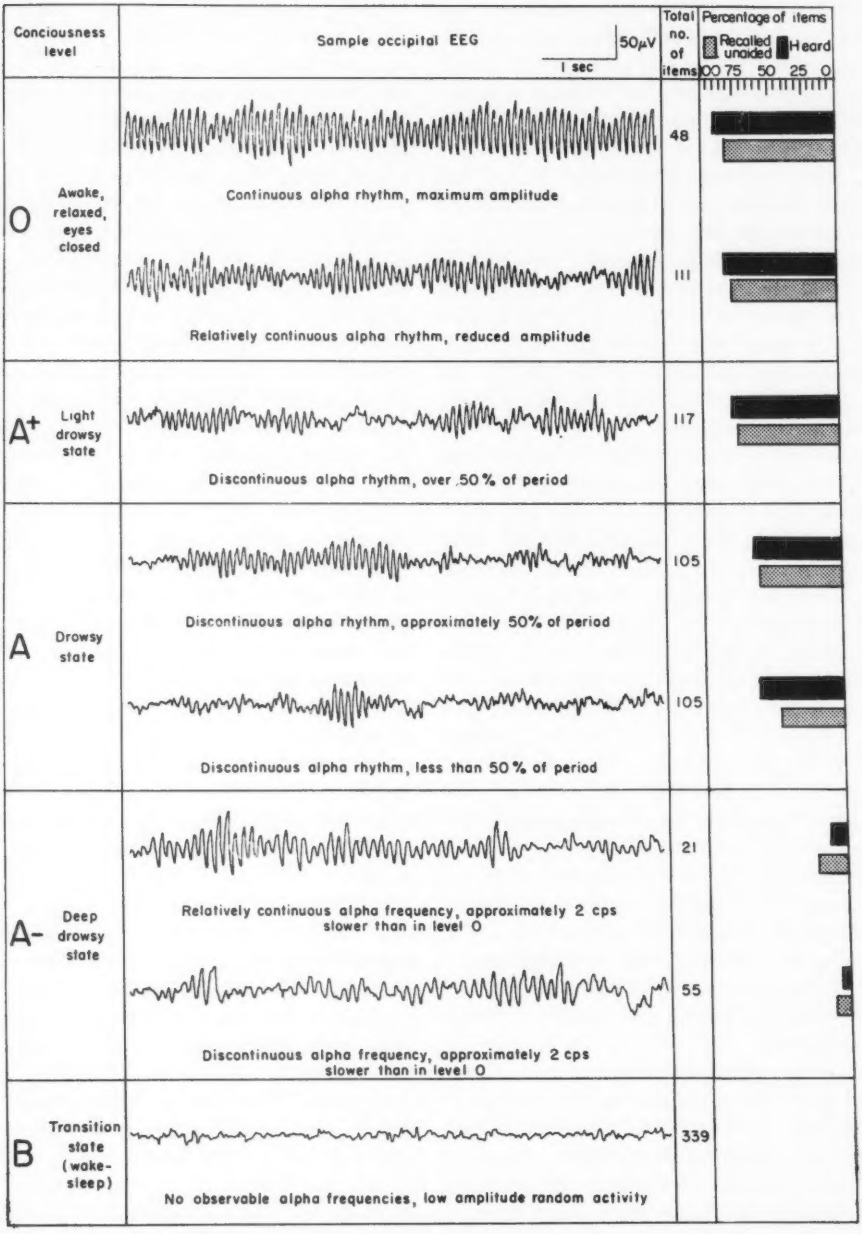


Fig. 2. Percentage of items heard and recalled during the appearance of waking electroencephalograms. (Reprinted from *Science* (Ref. 31) by permission from the American Association for the Advancement of Science.)

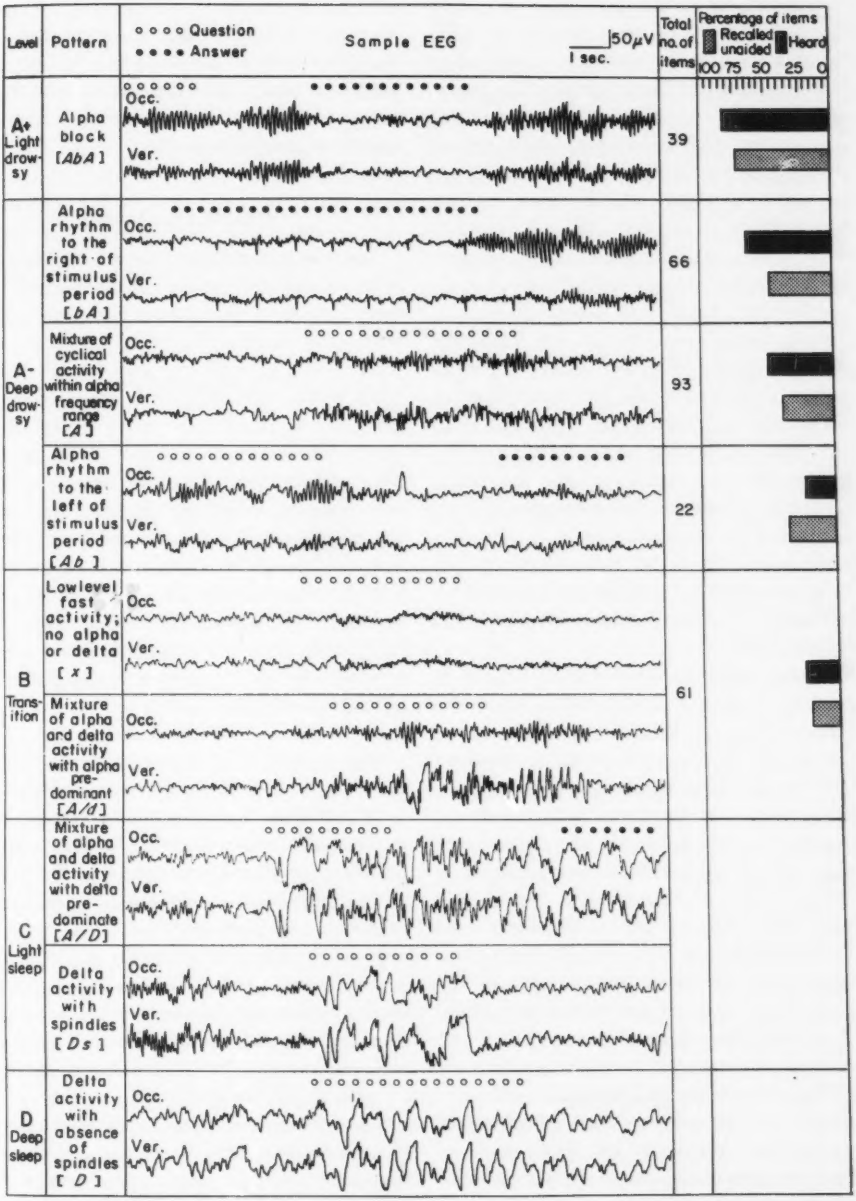


Fig. 3. Percentage of items heard and recalled during the appearance of electroencephalographic patterns modified by stimulation. (Reprinted from *Science* (Ref. 31) by permission from the American Association for the Advancement of Science.)

K-complexes and mixed alpha and delta activity. The recall of items when these EEG patterns were present, however, continued to follow the rules previously stated. The probability of items being remembered was related to the quality and quantity of alpha present within plus or minus 30 sec of stimulation, and when alpha and delta frequencies were mixed, the probability of recall was related to whichever of these two frequencies was dominant. When activation patterns were observed in the transition state, some items were recalled.

Inertia effect of sleep. Simon and Emmons[31] also examined the effect on memory of items associated with a particular EEG pattern when a 60-sec period 5 min earlier contained alpha frequencies (i.e. the subject was awake to some extent) or did not contain alpha (i.e. the subject was asleep). It was found that when items were remembered at all, significantly fewer were recalled when alpha had been absent 5 min before than when it had been present. This depressant after-effect was labeled the "inertia effect of sleep".

Summary. As with the other measures of consciousness, memory is dependent upon the presence of some alpha activity (when it can be observed) and absence of delta activity. Memory for internally or externally generated stimuli was, in general, directly related to the amplitude, proportion, and frequency of the alpha waves (which when reduced are indicators of a depressed state of consciousness) and was absent with the occurrence of dominant delta waves, characteristic of sleep. Although these patterns were modified by blocking in the waking states and evoked potentials in the drowsy and sleep states, memory of items associated with certain EEG patterns continued to follow the same general principles. Memory in addition to being influenced by the level of consciousness designated by EEG patterns at the time of learning was also affected by the prior neuro-physiological states of the subject.

Although drowsiness appears to decrease the quantity of learning, it may be practical to utilize these depressed levels of consciousness for certain types of training. There are also some indications that during drowsy states, the influencing of attitudes and the solving of problems may be facilitated.[32] Evidence to support these views at this time is primarily anecdotal and inferential.

CONCLUSIONS

The studies described here have related EEG patterns to behavior and to the physiology of normal humans at various levels along the wake-sleep scale. The fact that these studies were done by different people, in different laboratories, with different equipment, and for different purposes has made it difficult to classify them into a nice clean matrix (or multidimensional space) encompassing consciousness, sleep, and the EEG. Much of the work is a non-systematic probing at the threshold of a new field; much of it lacks definitiveness and is primarily descriptive. The quality of the research varies considerably. And yet, in spite of all that, there exists an astounding amount of similarity and coherence in the data. Out of the noise the facts stand out strong, begging for refinement in future research.

The most interesting and the most consistent fact that occurs throughout the studies reported here was that behavioral manifestations of consciousness disappear concomitant with the presence of slow waves (designated "delta") and in subjects where alpha could be readily detected, consciousness diminished with the reduction of amplitude of the alpha rhythm and with a slowing of its frequencies. Whatever mechanisms operate to produce these characteristic patterns are obviously related in some fashion to the mechanisms responsible for conscious activity in man.

Many of the practical applications of this kind of information were discussed in the introduction to this paper. Before we can make optimum use of our basic knowledge on the subject and if we hope to improve the validity of our theories on what underlies this important segment of behavior, future research must

become more systematic in its approach. The topic of the effects of drowsiness and sleep on performance is a truly interdisciplinary one, and this must be reflected in the research. Problems must be examined from the system point of view, i.e. concern with the environment, the task, the organism, and the response. The complexity of the human requires that studies be complex, i.e. multi-dimensional. Both the conditions and the results must be quantized. Our measurements and our measuring techniques must be refined.

APPENDIX A
THE EEG AS A CRITERION OF SLEEP

Throughout this paper, electroencephalographic activity was used as a reference against which the various performance data were ordered. Its relation to performance during wakefulness, drowsiness, and sleep has been highly consistent. Just as our non-scientific judgment of whether a person is awake or asleep is subjective, so too is the use of the EEG in that it has become a useful tool only because it correlates highly with our less objective, introspective, and observational data. As a working hypothesis, the EEG has proven to be an acceptable criterion of wakefulness and sleep in the normal adult (Emmons and Simon[14]).

It is of interest to briefly expand this relationship to other circumstances and behaviors related to depressed levels of consciousness, i.e. subacon states. Thus, depressed and intermittent alpha rhythms characteristic of the drowsy state have been observed in: hypnotized subjects, isolated subjects, sleep deprived subjects, schizoid-type psychotics, subjects going to sleep, or just after awakening—all of which tend to show poorer capabilities to detect, discriminate, and/or remember. Slower and mixed alpha frequencies characteristic of the deeper drowsy state have also been observed in: dreamers, stimuli deprived subjects experiencing hallucinations, catatonics, older persons, and children—all of whom show a lowered responsiveness to the environment. The very slow delta waves characteristic of sleep are also found in some cases of coma, brain damage and disease. It is too premature to suggest what the underlying relation is and it is certain that the EEG's are merely concomitant effects along with the behaviors, but the consistency of the over-all patterns cries out for a broader investigation of the interrelation among these kinds of conditions.

There have been numerous attempts to find some good criteria of levels of wakefulness, drowsiness, and sleep. The behavioral criteria (such as movement, arousal thresholds and the like) have been discussed. Some autonomic measures which are modified with the onset of sleep have also been mentioned. But the EEG, in addition to giving evidence of being both valid and reliable when properly interpreted has other advantages.

1. Recording the EEG does not alter the sleep state being studied. (This is not so with many of the behavioral measures which require a response leading to arousal.)
2. It appears to be concomitant with levels of consciousness and sleep and less affected by completely irrelevant conditions. (This is not so with such measures as heart rate, intercranial blood volume, and blood pressure; they appear to be more related to the horizontal resting position rather than degrees of consciousness.)
3. The EEG gives a changing scale covering the differentiating the entire range from wakefulness to deep sleep. (This is not so with body temperatures and heart rate which are highly related to the metabolic rate during wakefulness, but significantly less so and thus unreliable during sleep.)
4. The EEG gives immediate knowledge of the sleep level at the time the record is being made. (This is not true with criteria which represent averages over time such as heart rate, number of movements, or certain autonomic measures such as blood pressure which have latency periods.)

5. In sleep, the EEG minimizes individual differences such that records can be immediately interpreted. (Although familiarity with a subject's brain waves is helpful, it is still possible to look at the EEG record of an otherwise normal subject and know with a high degree of certainty if he is awake or asleep. This is not the case with measures such as heart rate, respiration, blood pressure, and the like which all require individual baselines.)

There are some disadvantages to the EEG which might be mentioned, although these are in part due to our inadequate knowledge of the subject.

1. Not all subjects have good alpha rhythms. With beta-dominant subjects, it is considerably more difficult to distinguish the transition between levels A and B, although after sleep, their records are essentially identical.
2. It is difficult to study subjects who are moving about with their eyes open. Alpha is generally not present and the movements can create artifacts. Confining the individual for studies of sleep is acceptable, but for finer evaluations of conscious behavior these difficulties are limiting.
3. The EEG is a temporary indicator of stimulation as well as a measure of the general neurophysiological condition of the organism. This dual role is only serious until we better understand the mechanism of the EEG.
4. The EEG requires expensive and delicate equipment and some technical training to interpret. The information obtained will vary to some extent with electrode placement, how they are attached, and the physical adjustments of the equipment.
   (a) Implanted electrodes can give different patterns to stimulation than surface electrodes.
   (b) Multiple electrodes will differentially show the presence or absence of certain waves in different portions of the head.
   (c) The use of filters and other settings on the EEG equipment can change the apparent patterns, accenting some and rejecting others.

Different methods for analyzing data can provide different information. These are discussed in Werre[33] and in Supplement No. 4, *Electroenceph. Clin. Neurophysiol.* (1953).

REFERENCES

1. Armington, J. D. and Mitnick, L. L., The electroencephalogram and sleep deprivation. *J. Appl. Physiol.* **14,** 247–250, 1959.
2. Aserinsky, E. and Kleitman, N., Regularly occurring periods of eye motility, and concomitant phenomena during sleep. *Science,* **118,** 273–274, 1953.
3. Bjerner, B., Alpha depression and lowered pulse rate during delayed actions in a serial reaction test: A study of sleep deprivation. *Acta Physiol. Scand.* **19,** Suppl. 65, pp. 1–93, 1949.
4. Blake, H. and Gerard, R. W., Brain potentials during sleep. *Amer. J. Physiol.* **119,** 692–703, 1937.
5. Blake, H., Gerard, R. W. and Kleitman, N., Factors influencing brain potentials during sleep. *J. Neurophysiol.* **2,** 48–60, 1939.
6. *Curr. List Med. Lit.,* Vols. **27** to **36.** Armed Force Medical Library, Washington 25, D.C.
7. Davis, H., Davis, Pauline A., Loomis, A. L., Harvey, E. N. and Hobart, G., Human brain potentials during the onset of sleep. *J. Neurophysiol.* **1,** 24–38, 1938.
8. Davis, H., Davis, Pauline A., Loomis, A. L., Harvey, E. N. and Hobart, G., Electrical reactions of the human brain to auditory simulation during sleep. *J. Neurophysiol.* **2,** 500–514, 1939.
9. Davis, Pauline A., Effects of acoustic stimuli on the waking human brain. *J. Neurophysiol.* **2,** 494–499, 1939.
10. Dement, W. and Kleitman, N., Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroenceph. Clin. Neurophysiol.* **9,** 673–690, 1957.
11. Dement, W. and Kleitman, N., The relation of eye movements during sleep to dream activity: an objective method for the study of dreaming. *J. Exp. Psychol.* **53,** 339–346, 1957.
12. Durup, G. and Fessard, A., L'electrencephalogramme de l'homme. Observations psycho-physiologiques relatives a l'action des stimuli visuels et auditifs. *Année Psychol.* **36,** 1–35, 1935.
13. Emmons, W. H. and Simon, C. W., The non-recall of material presented during sleep. *Amer. J. Psychol.* **69,** 76–81, 1956.
14. Emmons, W. H. and Simon, C. W., Electroenceph. criterion of sleep. *ERDL-Tulane Symposium on Flicker* (L. M. N. Bach, Editor). New Orleans, La. Sponsored by the Dept. of Physiology, Tulane Univ., School of Medicine and the U.S. Army Engineering R. and D. Labs. pp. 183–194, April 6, 1957.
15. Hernandez-Peon, R. and Donoso, M., Influence of attention and suggestion upon subcortical evoked electrical activity. *Electroenceph. Clin. Neurophysiol.* **9,** 385–396, 1957.

16. HOAGLAND, H., Studies of brain metabolism and electrical activity in relation to adrenocortical physiology. Recent progress in hormone research. (Pincus, G., Editor) *Recent Progress in hormone research*. Academic Press, New York. Pp. 29–63, 1954.

17. HOFFMAN, B. F., SUCKLING, E. E. and BROOKS, C. M. *et al.*, Quantitative evaluation of sleep. *J. Appl. Physiol.* **8,** 361–368, 1956.

18. JASPER, H. H., CRUIKSHANK, R. M. and HOWARD, H., Action currents from the occipital region of the brain in man as affected by variables of attention and external stimulation. *Psychol. Bull.* **32,** 565, 1935.

19. JASPER, H. and SHADGRASS, C., Conditioning the occipital alpha rhythm in man. *J. Exp. Psychol.* **28,** 373–388, 1941.

20. JOUVET, M. and COURJON, J., Variation of the subcortical visual responses during attention in man. *Electroenceph. Clin. Neurophysiol.* **10,** 344, 1958 (Abstract).

21. KRASNOGORSKI, N. I., Bedingte und unbedinte Reflexe im Kindesalter und ihre Bedeutung fur die Klinik. *Ergeb. Inn. Med. Kinderheilk.* **39,** 613–730, 1931.

22. LI, C. L., JASPER, H. and HENDERSON, L., Jr., The effect of arousal mechanisms on various forms of abnormality in the electroencephalogram. *Electroenceph. Clin. Neurophysiol.* **4,** 513–526, 1952.

23. LOOMIS, A. L., HARVEY, E. N. and HOBART, G., Cerebral states during sleep, as studied by human brain potentials. *J. Exp. Psychol.* **21,** 127–144, 1937.

24. LOOMIS, A. L., HARVEY, E. N. and HOBART, G., Distribution of disturbance patterns in the human electroencephalogram with special reference to sleep. *J. Neurophysiol.* **1,** 413–430, 1938.

25. MILERIAN, E. A., Electrical activity of the cerebral cortex during attention to sound stimuli. *Voprosy Psikhologii*, **2,** 101–112, 1955. Reprinted in *The central nervous system. Translations from the Russian medical literature*. Bethesda, Md.: U.S. National Institute of Health, pp. 553–567, 1959.

26. MORGAN, C. T. and STELLAR, E., *Physiological Psychology*. (2nd Ed.) McGraw-Hill, New York, 1950.

27. ROWLAND, V., Differential electroencephalographic responses to conditioned auditory stimuli in arousal from sleep. *Electroenceph. Clin. Neurophysiol.* **9,** 585–594, 1957.

28. RUSINOV, V. S. and RABINOVICH, M. Y., Electroencephalographic researches in the laboratories and clinics of the Soviet Union. *Electroenceph. Clin. Neurophysiol.* Supplement 8, May 1958.

29. SIMON, C. W. and EMMONS, W. H., Learning during sleep? *Psychol. Bull.* **52,** 328–342, 1955.

30. SIMON, C. W. and EMMONS, W. H., Responses to material presented during various levels of sleep. *J. Exp. Psychol.* **51,** 89–97, 1956.

31. SIMON, C. W. and EMMONS, W. H., EEG, consciousness and sleep. *Science*, **124,** 1066–1069, 1956.

32. SIMON, C. W., Some applications of repetitive verbal stimulation during drowsy states. 1957. (Unpublished paper.)

33. WERRE, P. F., The relationships between electroencephalographic and psychological data in normal adults. Univ. Leyden Press, Leyden, Holland, 1957.

34. WILLIAMS, H. L., LUBIN, A. and GOODNOW, JACQUELINE., Impaired performance with acute sleep loss. *Psychol. Monogr.* **73,** No. 14 (Whole No. 484), 1959.

A Technique for Evaluating Man–Machine System Designs[1]

ARTHUR I. SIEGEL and J. JAY WOLF, Applied Psychological Services, Wayne, Pa.

A computer based method for digitally simulating the performance, in one-operator systems, of operators who possess various characteristics is described. The method is believed applicable for evaluating various system designs while the system is in the early design stage. Two operational tasks, landing an F4D aircraft on a carrier and firing an air-to-air missile, were simulated using the method. The predictions from the model were compared with outside criterion data for the same tasks. The predictions are held to conform generally with reality and to be reasonable. The results of the two applications of the model were in general agreement. It is held that the model may be considered sound and may now be tentatively employed for comparative evaluation of alternative system designs or for predicting system performance.

INTRODUCTION AND DESCRIPTION OF MODEL

Complex systems in which a human is expected to operate or control his machine are often found to overburden or underburden the operator after the system is accomplished. In order to achieve the general goal of simulating the characteristics of the human operator in a man–machine system, a psychological "model" or method has been developed which allows a determination of where a single-operator man–machine system may overload or underload the operator while the system is in the design stage.

It is the purpose of this model to give the equipment designers quantitative answers to questions such as the following while the equipment is in the early design stage:

(1) Can an average operator be expected to complete successfully all actions required in the performance of a task within a time limit, T, for a given operator procedure and a given machine design?
(2) How does success probability change for slower or faster operators and longer or shorter periods of allotted time?
(3) How great a stress is placed on the operator during his performance and in which portions of the task is he overloaded or underloaded?
(4) What is the frequency distribution of operator failures as a function of various stress tolerances and operator speeds?

Use of the model is based on the high-speed, general-purpose digital computer. The computer operates on source data and simulates an operator by calculating values for, and keeping track of, items such as the operator's response time, failures, stress, and idle time for each action (called a subtask) which the operator must perform. The speed of arithmetic computation, capacity for storing information and the flexibility of decision making offered by such computing devices enabled the computer, in approximately 40 sec, to simulate an entire landing operation which must normally be completed by the pilot in 210 sec. This task which consists of from 37 to 41 subtasks was simulated almost 3000 times by the computer representing 37 pilot "types".

Similarly, a missile launching task which must be completed by the pilot in at most 184.8 sec, was simulated in approximately 40 sec of computation. This task, consisting of from 22 to 24 subtasks, was simulated 3080 times representing 35 pilot "types". This paper describes the model's rationale, method of development, and the results obtained to date from its applications.

Overview of Model

Prior to the use of the model a task analysis is performed for the man–machine task under

[1] This program was carried out under Contract Nonr-2492(00) between Applied Psychological Services and the Office of Naval Research.

consideration. For each subtask certain specific required source data are compiled. These data together with the parameter values selected are prepared in punched card form and introduced into the digital computer for which a computer program has been previously prepared. As directed by its program the computer will simulate the performance of each subtask by the operator in the proper sequence, according to the rules of the model. The simulation, consisting of calculating stress, urgency, execution time, the determination of success and the proper book-keeping continues serially for each subtask. A simulation is completed when the simulated operator either runs out of allotted time or successfully completes the task. Following the computer's "performance" of the task, a set of output cards is punched by the computer and their data are printed to indicate the areas of operator overload, failures, idle times, urgency conditions, etc., for the given set of selected parameters. Numerous repetitions of the task with different parameter values yield additional punched cards and printouts. Frequency distributions and graphs are prepared either automatically from the cards or manually from the printed records. If alternative designs are indicated by analysis of the results the new designs are similarly prepared and analyzed in order to determine the extent of improvement brought about by the alternative designs.

For each subtask i ($i = 1, 2, \ldots, n$), the model considers and involves the following basic data as derived from the best available sources:

(1) the average time, $\bar{t}_i$, required by the operator to perform the subtask. This represents the operator's average execution time when he is under no stress;
(2) the average standard deviation, $\bar{\sigma}_i$, around the mean, $\bar{t}_i$, for an average operator (no stress);
(3) the average probability, $\bar{p}_i$, of performing the subtask successfully (no stress);
(4) an indication of necessity, i.e. whether or not the successful performance of the subtask is essential to successful completion of the task;
(5) a waiting time (if applicable), W_1, i.e. the number of seconds after starting before which the operator cannot begin the subtask;
(6) the subtask to be performed next in the event of failure of subtask i;
(7) the subtask to be performed next in the event of success of subtask i.

Using these data, the computer executes three principle calculations for each subtask in turn. First, the urgency and stress conditions are computed as a function of the work remaining for the operator and the time available.

Conditions for the three states of urgency considered are:

(1) The situation is non-urgent when, on the average, sufficient time remains to complete all remaining subtasks assuming average speed and no operator failures. Thus, the non-urgent state exists during performance of subtask i when:

$$\sum_{j=i}^{n} \bar{t}_j \leqslant \text{Time remaining to complete task} \leqslant T - T_i^U$$

where T is the total time allowed for task completion and T_i^U is the total of all time used in performing all subtasks through i minus 1 (including required repetitions, if any).

(2) The urgent state occurs whenever there is insufficient time available to complete all remaining subtasks, provided that sufficient time is available to complete all remaining essential subtasks. That is, the situation is urgent if:

$$\bar{T}_i^E + \bar{T}_i^N = \sum_{j=i}^{n} \bar{t}_j > T - T_i^U$$

and

$$\bar{T}_i^E \leqslant T - T_i^U$$

where $\bar{T}_i^E$ and $\bar{T}_i^N$ are the total times to complete the remaining essential and non-essential subtasks respectively including the ith subtask. These are based on estimated average operator execution

TABLE I

Summary of Urgency and Stress Conditions

Urgency	$\frac{\bar{T}_i^E + \bar{T}_i^N}{T - T_i^U}$	$\frac{\bar{T}_i^E}{T - T_i^U}$	*Condition (assume average operator)*	*Result*	*Stress*
Non-urgent	$\leqslant 1$	$\leqslant 1$	Time is available to perform all remaining subtasks	Perform all subtasks	1
Urgent	> 1	$\leqslant 1$	Time is available to perform all remaining essential subtasks but not all remaining subtasks	Ignore non-essential subtasks	1
Highly urgent	> 1	> 1	Time is not available to perform all essential subtasks	Ignore non-essential subtasks; compute stress	$\frac{\bar{T}_i^E}{T - T_i^U}$

times and the assumption that no operator failures will occur. In this situation the non-essential tasks are ignored by the computer.

(3) the situation is highly urgent if there is insufficient time available for completing even the remaining essential subtasks at average operator speeds.

The "certainty" in the operator's "mind" that there is insufficient time remaining to complete the essential subtasks (when performing at normal speed and efficiency) will impress a state of stress on the operator. The model defines stress, s_i, as a state of mind of the operator just prior to his performance of subtask i. Theory suggests that emotion or stress up to a certain point acts as an organizing agent on behavior; beyond this point stress acts as a disorganizing agent. Accordingly, the model recognizes an organizing effect on operator performance as long as s_i is less than M, where M is a threshold value; if s_i equals or exceeds M, the effect is disorganizing. During non-urgent and urgent conditions, $s_i = 1$, and when the situation is highly urgent, stress is defined as the ratio of the sum of the average remaining essential subtask execution times to the total time remaining:

$$s_i = \frac{\bar{T}_i^E}{T - T_i^U} \geqslant 1$$

Thus, stress is the ratio of how much is left to do to the amount of time available in which to do it. The stress threshold, M, may be considered as the operator's breaking point. For example, an M value of 2.0 indicates that the operator begins to become slower and less accurate at the point at which he has more than twice as much to do (at average speed) as he has time available. Prior to this point, the added backlog of essential subtasks creates a mental inducement of stress which makes his actions faster and more accurate. Table I summarizes the conditions for urgency and stress as incorporated into the model.

Second, the execution time of the subtask is computed. The average operator will require $\bar{t}_i$ seconds to perform subtask i when s_i equals unity. In this case his average standard deviation will be $\bar{\sigma}_i$. Of course, no two operators would be expected to perform any subtask in exactly the same time, nor would any one operator complete a subtask in exactly the same time on each repetition. For each subtask, however, it is assumed that the actual subtask execution time, t_i, for a specific subtask, i, is normally distributed with mean dependent on $\bar{t}_i$ and standard deviation dependent on $\bar{\sigma}_i$. The computation of a reasonable and realistic specific value for the actual execution time, t_i, on each subtask is made on the basis of a random selection from a normal distribution

limited from below by a fixed minimum. This minimum is the average time (selected as 1.5 sec) required to decide to skip or ignore a subtask. The computer accomplishes the selection of t_i by a random or Monte Carlo technique. Pseudo-random numbers uniformly distributed in the unit interval are sequentially generated by the computer using Lehmer's method as described by Stibitz and Larrigee.[2] By the use of these independent random numbers corresponding values of an independent random variable are generated with a distribution function equivalent to that of the normal distribution. This is done by the rejection technique described by Kahn.[1]

Thus, if K_i is the number of standard deviations from the mean corresponding to the random numbers generated in simulating subtest i, then the actual simulated execution time utilized by the computer, t_i, is given by:

$$t_i = \begin{cases} \left[\bar{t}_i \pm \dfrac{K_i \bar{\sigma}_i}{s_i}\right] F, & \text{if } s_i < M \\ [\bar{t}_i (s_i + 1 - M) \pm K_i \bar{\sigma}_i (2s_i + 1 - 2M)]\, F, & \text{if } M \leqslant s_i \leqslant M+1 \\ [2\bar{t}_i \pm 3K_i \bar{\sigma}_i]\, F, & \text{if } s_i > M+1 \end{cases}$$

where F is an individuality factor described later. The effect of the above is to provide a t_i value in which the values of $\bar{t}_i$ and $\bar{\sigma}_i$:

(1) are used unchanged when stress equals unity;

(2) are decreased with increasing stress until stress assumes the threshold value;

(3) are used unchanged when stress equals the threshold value;

(4) are increased linearly with increasing stress beyond M until, when stress equals $M+1$, the contributions of $\bar{t}_i$ and $\bar{\sigma}_i$ remain constant at $2\bar{t}_i$ and $3\bar{\sigma}_i$ respectively.

Third, the computer determines whether or not to consider the subtask as performed successfully by the operator. The average probability of successful performance of subtask i, $\bar{p}_i$, is given as an input for each subtask. However, the actual probability of successful performance of a given subtask, p_i, is a function of $\bar{p}_i$, stress, s_i, and the stress threshold, M, as follows:

$$p_i = \begin{cases} \bar{p}_i + \dfrac{(1-\bar{p}_i)(s_i - 1)}{M-1}, & \text{if } s_i < M \\ \bar{p}_i (s_i + 1 - M) + (M - s_i), & \text{if } M \leqslant s_i \leqslant M+1 \\ 2\bar{p}_i - 1, & \text{if } s_i > M+1 \end{cases}$$

Thus, probability of success increases linearly with stress until it reaches certainty at the threshold of stress. At this point, the probability assumes the average value, $\bar{p}_i$, after which it decreases linearly until, when stress has a value equal to $M+1$, it assumes a constant value which is decreased from $\bar{p}_i$ by an amount $1-\bar{p}_i$. In order to determine actual success or failure for any subtask, the computer generates a pseudo-random number, R, according to the method previously referenced. The subtask is considered to be performed successfully if R is less than p_i; otherwise it is assumed that the operator failed to perform the subtask properly.

Many subtasks require more than one control action. In this case, subtask success is defined as the completion of all actions; the average execution time for the subtask is taken as the product of the average time required for a single action times the number of actions; the standard deviation is treated similarly. The probability of success is associated with the combined multiple operation. Certain other cases (called multiple action subtasks) arise in which repetitious trials of the same action are usually required in order to gain a given probability of success, but in which a single successful action at any time is sufficient for subtask success. In these cases, although it is improbable to succeed on the first trial, it is possible. Hence, the model treats these subtasks as one subtask which requires a single control action but which possesses a low success probability. The actual determination

of the average probability of success on any single trial, p_i, in multiple action subtasks is:

$$p_i = 1 - \sqrt[n]{(1-p^*)}$$

where p^* is the probability of at least one success after n trials. This method produces failure and success indications according to these computed probabilities in the long run. In the event of either success or failure, input information indicates the subtask to be performed next.

Initial Conditions and Parameters

Since the simulation of any individual subtask is based in part on a random effect, it is necessary to repeat the simulation of a task many times. The number of times, N, that a given task is to be simulated by the computer is the first initial condition inserted into the computer prior to a computer run for a given task. Thus, there are N simulations per computer run. Another initial condition is R_0, the number from which the computer generates subsequent pseudo-random numbers during the course of the calculations. The initial condition value, R_0, selected is 12345678 for the first run performed. Subsequently, the last pseudo-random number generated is used as the first value in the next run.

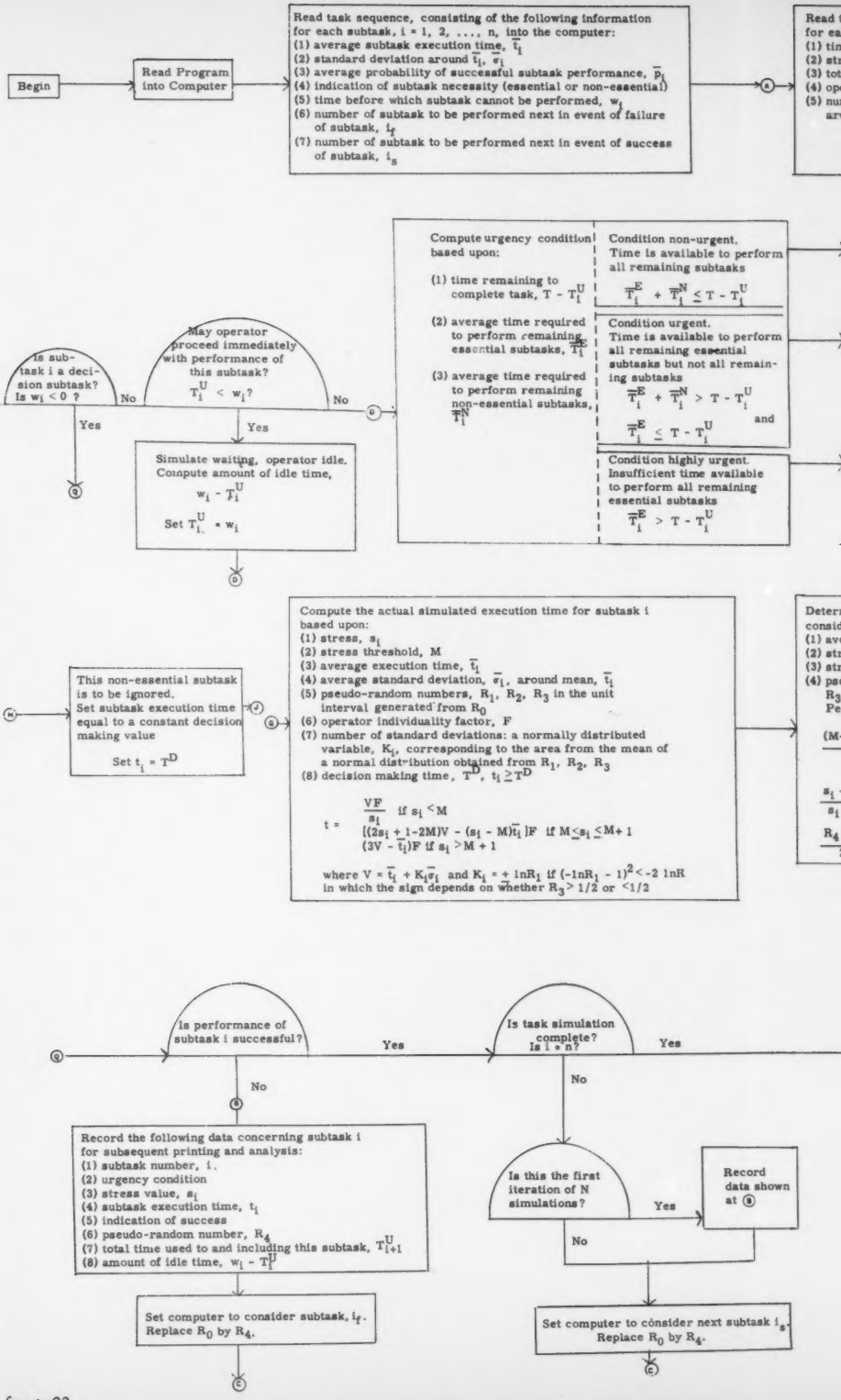
Parameters are those initial conditions, selected prior to a computation, which one may want to vary in order to evaluate either the model or a man–machine system. The stress threshold, M, is such a parameter. For any given run of N simulations, a specific stress threshold is used in the calculation. The effect of a change in M is studied by performing runs with different values of M. A second parameter, T, is the total time allotted to perform the task. Various computer runs may be performed to determine the effect of a change of this parameter within reasonable limits. This facility is provided since the value may not be known exactly in systems which are proposed or designed but not yet built.

The parameter F which accounts for variance among individuals is an individuality factor. This provides the ability to simulate an operator who usually performs faster or slower than, or whose interest differs from, the average. The parameter F possesses a value of unity for an average operator. The effect of faster or more highly motivated operators ($F < 1$) and slower operators ($F > 1$) in the performance of the task is examined by performing several computer runs with different values for this multiplicative factor. Thus for a given run, N simulations of a single task are performed by the computer with a fixed value for each of the parameters M, T, F, starting with R_0 as the first pseudo-random number.

A Review of Calculations

Fig. 1 is a general flow chart of the model. The computer accepts the program, data (consisting of information concerning the task sequence), initial conditions, and parameters. Once these are entered via punched cards and stored simultaneously in the computer's memory the computer automatically performs runs of N simulations without human intervention. The prevailing urgency condition and state of stress, s_i, are determined. Using these, the computer calculates the subtask execution time, t_i, and p_i, from which it determines whether or not the subtask was performed successfully.

The amount of time remaining to the operator for task completion is then adjusted. On the basis of the computed probability (indicating either success or failure of the subtask), a decision is made as to which subtask to perform next. In the case of subtask failure the next task tried by the operator may consist of a repetition of the task last performed; if the previous task was successfully performed then the next subtask is considered by the computer. This process continues for each subtask considered in turn. A simulation is terminated when the operator either runs out of time or when he successfully completes the task. After each task simulation the computer causes pertinent information to be recorded for subsequent


Begin
Read Program into Computer
Read task sequence, consisting of the following information for each subtask, i = 1, 2, ..., n, into the computer:
(1) average subtask execution time, $\bar{t}_i$
(2) standard deviation around $\bar{t}_i$, σ_i
(3) average probability of successful subtask performance, $\bar{p}_i$
(4) indication of subtask necessity (essential or non-essential)
(5) time before which subtask cannot be performed, w_i
(6) number of subtask to be performed next in event of failure of subtask, i_f
(7) number of subtask to be performed next in event of success of subtask, i_s
Is sub-task i a decision subtask? Is $w_i < 0$?
No
Yes
May operator proceed immediately with performance of this subtask? $T_i^U < w_i$?
No
Yes
Simulate waiting, operator idle. Compute amount of idle time, $w_i - T_i^U$
Set $T_{i.}^U = w_i$
Compute urgency condition based upon:
(1) time remaining to complete task, $T - T_i^U$
(2) average time required to perform remaining essential subtasks, $\bar{T}_i^E$
(3) average time required to perform remaining non-essential subtasks, $\bar{T}_i^N$
Condition non-urgent. Time is available to perform all remaining subtasks
$\bar{T}_i^E + \bar{T}_i^N \leq T - T_i^U$
Condition urgent. Time is available to perform all remaining essential subtasks but not all remaining subtasks
$\bar{T}_i^E + \bar{T}_i^N > T - T_i^U$ and $\bar{T}_i^E \leq T - T_i^U$
Condition highly urgent. Insufficient time available to perform all remaining essential subtasks
$\bar{T}_i^E > T - T_i^U$
This non-essential subtask is to be ignored. Set subtask execution time equal to a constant decision making value
Set $t_i = T^D$
Compute the actual simulated execution time for subtask i based upon:
(1) stress, s_i
(2) stress threshold, M
(3) average execution time, $\bar{t}_i$
(4) average standard deviation, $\bar{\sigma}_i$, around mean, $\bar{t}_i$
(5) pseudo-random numbers, R_1, R_2, R_3 in the unit interval generated from R_0
(6) operator individuality factor, F
(7) number of standard deviations: a normally distributed variable, K_i, corresponding to the area from the mean of a normal distribution obtained from R_1, R_2, R_3
(8) decision making time, T^D, $t_i \geq T^D$
$t = \frac{VF}{s_i}$ if $s_i < M$
$[(2s_i + 1 - 2M)V - (s_i - M)\bar{t}_i]F$ if $M \leq s_i \leq M+1$
$(3V - \bar{t}_i)F$ if $s_i > M + 1$
where $V = \bar{t}_i + K_i\bar{\sigma}_i$ and $K_i = +\ln R_1$ if $(-\ln R_1 - 1)^2 < -2 \ln R$ in which the sign depends on whether $R_3 > 1/2$ or $< 1/2$
Is performance of subtask i successful?
Yes
No
Is task simulation complete? Is i = n?
Yes
No
Record the following data concerning subtask i for subsequent printing and analysis:
(1) subtask number, i.
(2) urgency condition
(3) stress value, s_i
(4) subtask execution time, t_i
(5) indication of success
(6) pseudo-random number, R_4
(7) total time used to and including this subtask, T_{i+1}^U
(8) amount of idle time, $w_i - T_i^U$
Is this the first iteration of N simulations?
Yes
No
Record data shown at (B)
Set computer to consider subtask, i_f. Replace R_0 by R_4.
Set computer to consider next subtask i_s. Replace R_0 by R_4.

To face p. 22

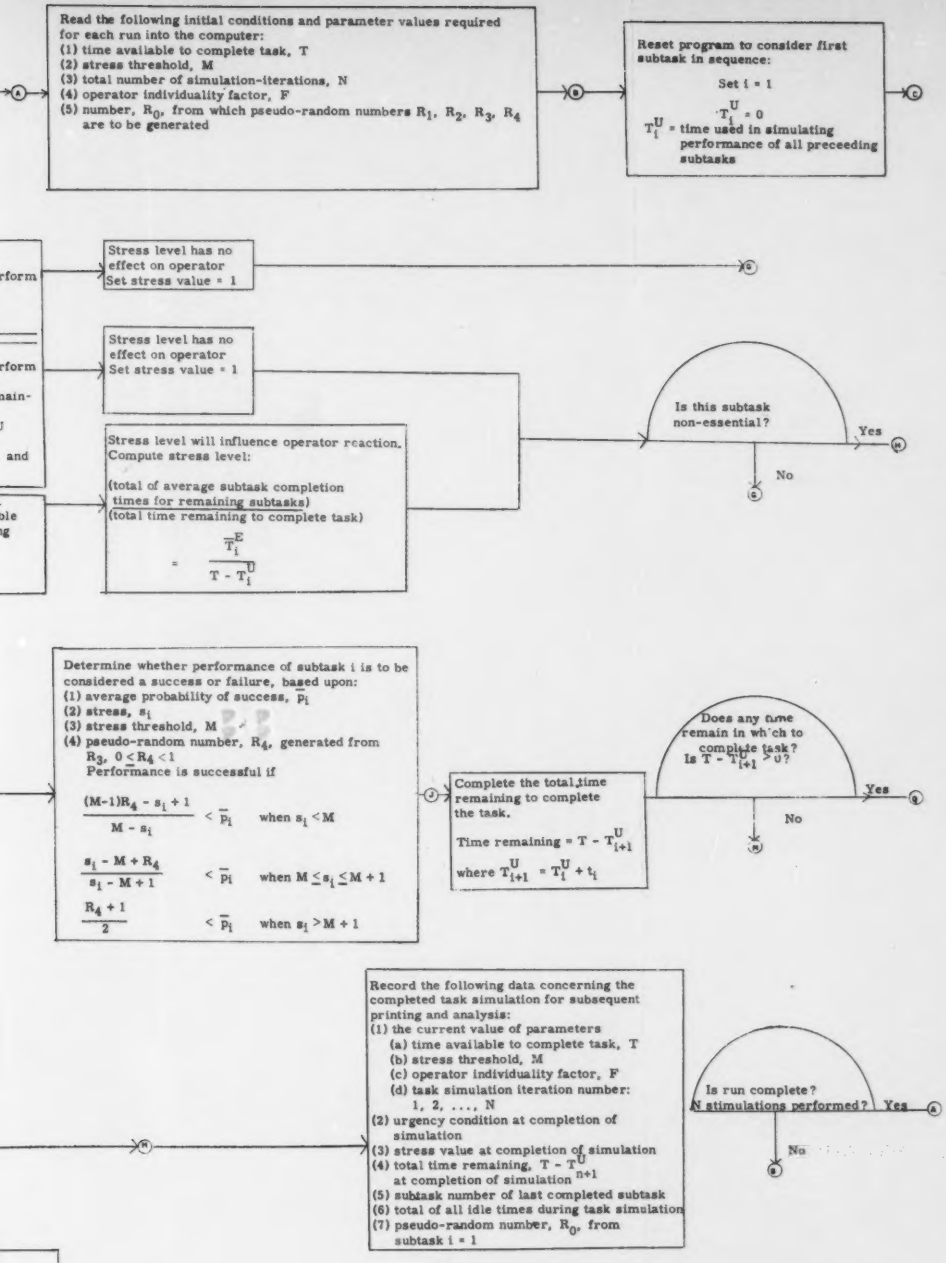

Read the following initial conditions and parameter values required for each run into the computer:
(1) time available to complete task, T
(2) stress threshold, M
(3) total number of simulation-iterations, N
(4) operator individuality factor, F
(5) number, R_0, from which pseudo-random numbers R_1, R_2, R_3, R_4 are to be generated
Reset program to consider first subtask in sequence:
Set i = 1
$T_i^U = 0$
T_i^U = time used in simulating performance of all preceeding subtasks
Stress level has no effect on operator
Set stress value = 1
Stress level has no effect on operator
Set stress value = 1
Stress level will influence operator reaction.
Compute stress level:
(total of average subtask completion times for remaining subtasks)
(total time remaining to complete task)
$= \frac{\overline{T}_i^E}{T - T_i^U}$
Is this subtask non-essential?
Yes
No
Determine whether performance of subtask i is to be considered a success or failure, based upon:
(1) average probability of success, $\overline{p}_i$
(2) stress, s_i
(3) stress threshold, M
(4) pseudo-random number, R_4, generated from R_3, $0 \leq R_4 < 1$
Performance is successful if
$\frac{(M-1)R_4 - s_i + 1}{M - s_i} < \overline{p}_i$ when $s_i < M$
$\frac{s_i - M + R_4}{s_i - M + 1} < \overline{p}_i$ when $M \leq s_i \leq M + 1$
$\frac{R_4 + 1}{2} < \overline{p}_i$ when $s_i > M + 1$
Complete the total time remaining to complete the task.
Time remaining = $T - T_{i+1}^U$
where $T_{i+1}^U = T_i^U + t_i$
Does any time remain in which to complete task?
Is $T - T_{i+1}^U > 0$?
Yes
No
Record the following data concerning the completed task simulation for subsequent printing and analysis:
(1) the current value of parameters
(a) time available to complete task, T
(b) stress threshold, M
(c) operator individuality factor, F
(d) task simulation iteration number: 1, 2, ..., N
(2) urgency condition at completion of simulation
(3) stress value at completion of simulation
(4) total time remaining, $T - T^{U}_{n+1}$ at completion of simulation
(5) subtask number of last completed subtask
(6) total of all idle times during task simulation
(7) pseudo-random number, R_0, from subtask i = 1
Is run complete?
N stimulations performed?
Yes
No

Fig. 1. General flow chart of the model.

analysis and reduction. Frequency distributions and graphs are then prepared either automatically from the resultant punched cards or manually using records printed therefrom.

Provision for skipping one or more subtasks or selecting an alternate sequence of subtasks with a predetermined probability is incorporated by a decision subtask feature in which the simulated operator's first action is to decide on the subtask to perform next.

Decision subtasks may be placed anywhere in the sequence of subtasks desired. The ith subtask is considered a decision subtask when W_i is negative. In this case the values of $\bar{t}_i$, $\bar{\sigma}_i$, and essentiality have no meaning. When a negative waiting time is detected by the computer, the t_i calculation is by-passed and the last pseudo-random number is compared against p_i of the decision subtask. Therefore the next subtask to be performed is the one normally performed next in the event of success, with probability $\bar{p}_i$, and the subtask indicated to be performed in the event of failure is executed with probability $1-\bar{p}_i$. Thus both branching, skipping and looping are made possible.

The computer program is a list of 616 basic coded instructions and constants in the form of a deck of 141 punched cards which direct the sequence of computer operations during the simulation. The program accommodates a task sequence of 100 subtasks (including decision subtasks).

Assumptions

It is assumed in the use of the model that the operator remembers and executes the correct sequence of subtasks. It should be noted, however, that the possibility of an operator's neglecting a subtask or of rearranging the performance of a sequence of subtasks may be studied by additional runs using these different sequences, i.e. assuming new tasks.

Similarly, a change in the predetermined sequence of subtasks is conceivable in the event of emergency. Such a dangerous situation may result from operator action or an external event during the task. In either case it may be assumed that the operator, upon noticing the danger will abandon his task and take up the problem of survival, and the sequence of operations will change. Thus, a dangerous condition need not be a part of the model since the danger condition may itself be studied using the model by establishing special danger sequences to be simulated. Therefore, this situation does not limit the model.

It must also be assumed that $\bar{p}_i$, $\bar{t}_i$ and $\bar{\sigma}_i$ are independent of whether the subtask is being performed for the first time or is being repeated due to a previous failure. However, it is noted that upon repetition, less time will remain for the operator, possibly affecting the stress and consequently affecting p_i, t_i, and σ_i.

Application of Model

The model as described in the preceding sections was applied to two single-operator man–machine tasks: landing an F4D aircraft aboard an aircraft carrier and launching a SPARROW II air-to-air missile. Since the purpose of these applications was validation of the model, the "predictions" from the model were compared with independently derived, outside criteria data. First, the landing task was simulated; then the launching task was analyzed. Thus, it was possible to isolate M values in the landing task and cross validate these values in the launching task.

Task Analysis—Landing

A task analysis indicated that the landing task is composed of 37 basic subtasks from the start of the maneuver to the last operator action required in the final approach. After subtask number 37 a decision with the following five possible first-order alternatives is made by the pilot:

(1) the aircraft is approaching satisfactorily and no further operator action is required. In this case the simulation is complete;
(2) the aircraft is approaching too fast and subtasks 40 through 43 must be performed;

(3) the aircraft is approaching too slow and subtasks 45 through 48 must be performed;
(4) the aircraft is approaching too high and subtasks 50 through 53 must be performed;
(5) the aircraft is approaching too low and subtasks 54 through 57 must be performed.

Second-order alternatives, such as high and fast, were not considered. The assumption that these five listed alternatives were equally likely (i.e. each had a probability of 0.2) was implemented by the technique of decision subtasks. For each subtask of the landing sequence the subtask to be performed next in the event of a failure was the same subtask which failed. This meant that each subtask was repeated until it was performed correctly.

A total of 210 sec is available to the pilot during which the sequence of subtasks is to be performed, i.e. $T = 210$ for all runs of the landing task. Because of the nature of the landing task, the pilot cannot begin the execution of subtask 27 until 141 sec after the start of the landing operation. This usually leads to a time during which the plane may maneuver to the point where a "mirror" on the landing deck is visible to the pilot.

Task Analysis—Missile Launching

A task analysis for the launching task indicated 22, 23, or 24 basic subtasks from the start of the maneuver to the final "breakaway". Subtask 13, a decision subtask, allowed the computer during simulation to ignore subtask 14 25 per cent of the time. Following subtask 19, a decision is made in accordance with the following (at least) three equally probable alternatives:

(1) the target is properly centered on the radar scope; therefore subtasks 21 through 24 are to be performed;
(2) the target is too high on the radar scope; therefore subtasks 26 through 30 must be performed;
(3) the target is too low on the radar scope; therefore subtasks 31 through 36 must be performed.

Because of the nature of the launching task, subtask number 18 could not be initiated until "lock-on" had occurred. This was allowed after a minimum of 79.2 sec following the start of the maneuver.

The value of T was determined from the best estimate of the distance traveled by the intercept aircraft (including maneuvers) and its average speed. The value of T selected for runs 1 through 25 of the launching task was 184.8 sec, based on a breakaway distance of two miles between the attacker and his target. Runs 26 through 35 were made to determine the general magnitude of task completion if a 4-mile breakaway distance is employed. In this case, $T = 171.6$ sec. A simulation run was considered successful only if all subtasks, including breakaway, were completed prior to T elapsed seconds. In practice, if insufficient time remained to complete the required action prior to missile launch, the pilot would, of course, give up the launching sequence and breakaway.

Outside Criterion—Landing

In order to obtain some measure of the agreement between the results obtained from the model and the actual performance of pilots outside, criterion data were obtained. The outside criterion data, derived from ratings of Landing Signal Officers for 162 F4D carrier landings are presented in Table II.

Outside Criterion—Missile Launching

Outside criterion data for the launching task, fleet records, indicated that 713 missile launching attempts out of 876 were successful, an overall probability of task success of 81.4 per cent. Here success indicates a missile launching and does not necessarily imply a "kill".

Results, Landing Task

As may be expected, values of M and F were found to have a pronounced effect on the

TABLE II

Outside Criteria Data for the Carrier Landing Task

Class of landing	*Flight description*	*Frequency*	*Percentage*
1	Poor landing, given waveoff or resulting in bolter (bounce on deck)	41	25.3
2	Mediocre landing, improvement highly desirable	59	36.5
3	Good approach but improvement possible	41	25.3
4	Excellent landing	21	12.9
		162	100.0

number of subtasks which were completed by the simulated pilot. In the worst case encountered the task was failed by the operator after completion of as few as 27 subtasks. In the best case, no landing failures were encountered for an entire run. Frequency distributions were prepared of the last subtask performed for each run. The data were grouped into classes of landings to facilitate summarization and comparison with the outside criterion data. Class 1 was defined, in this case, as those landings which were not completed to the decision point following task 37, i.e. those simulations in which the last subtask to be successfully completed was subtask 36 or below. Class 2 included those simulations in which the decision point was reached, further adjustments were required, but these were not completed successfully before all allotted time was expended. Class 3 included those simulations in which the additional adjustments were required and completed satisfactorily. Class 4 included those simulations in which the decision point was reached successfully, no further adjustments were required, and the last subtask completed was 37. Thus, Class 4 corresponds to a perfect landing. Class 3 corresponds to a landing in which the pilot completed his job (including last second adjustments).

These data in the four classes were further aggregated into two divisions—those simulations in which the task was not completed (Classes 1 and 2), and those simulations in which it was (Classes 3 and 4). This grouping, plotted in Fig. 2 as a function of the stress threshold yields a value for the number of failures in each of the 37 computer runs.

If one accepts the tenuous assumptions that the aggregate of Classes 3 and 4 as given in Table II (good and excellent landings) corresponds to a simulation in which the pilot has completed all subtasks required (i.e. they are equivalent to Classes 3 and 4 of the computer simulation) and that the aggregate of Classes 1 and 2 as given in Table II (mediocre and poor landings) corresponds to the Classes 1 and 2 of the computer simulations, then the probability of failure from outside criteria data, obtained from Table II as 0.618 (50 out of 81 failures) may be compared to computed data. This value of 50 is indicated on Fig. 2 as an arrow at the Y axis at $M = 2.3$. Subject to the assumptions made above, the outside criteria data may be used now to indicate a range of values for the stress threshold, M.

Since a primary aim of this study was to evaluate the model and determine conditions under which it agrees with actual events, this value of $M = 2.3$ is a value of a basic parameter under which agreement is obtained between computed results and outside criteria. Realizing, however, that the individual points in Fig. 2 are indicative only within certain confidence limits, the M value should more

accurately be given by a range of values. This range of values taken from Fig. 2 is approximately $1.95 \leqslant M \leqslant 2.80$ corresponding to failure frequencies from 46.5 to 53.5 out of 81, or failure probabilities of from 0.575 to 0.66.

Another comparison may be made between the outside criterion data and the computed results if the same assumptions described earlier in this section are made. Specifically, tests for the significances of the differences between the outside criterion data and summaries of computed data may be performed using the Chi square test. This test indicated that a wide range of non-significant differences was achieved. Moreover, the entire area of M values (1.95–2.80) which were indicated as the most possible valid model parameters yielded non-significant differences when compared with the outside criterion.

The results of the simulations indicated that in every simulated landing some "idle time" was encountered by the pilot on the "downwind leg" prior to picking up the mirror (subtask 27).

These data seem to agree fairly well with pilot opinion which suggests that on making the approach the pilot is not overloaded on the "downwind leg" but that he is very busy during the "final approach".

In summary, consider the case of an average speed pilot ($F = 1$). Outside criterion data indicated a probability of 0.618 for a poor or mediocre landing. If we assume that these classes of landings correspond to computer simulations in which all the available time was consumed before all subtasks were completed, then we find that a task failure rate equal to these classes resulted from this model when the stress threshold had an interpolated value of 2.30. The accuracy of the computational method considering the number of simulations performed, implies that this value actually may

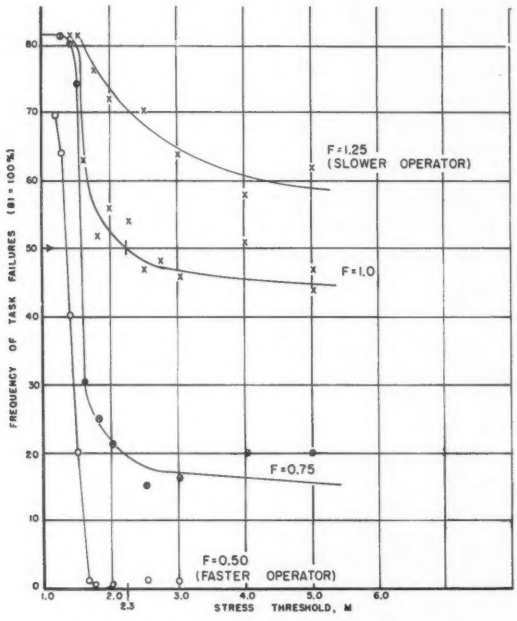


Fig. 2. Frequency of task failures, landing task.

encompass a range between 1.95 and 2.80. The results pertaining to task failures were not significantly different from outside criterion data over this entire range, which corresponds to task failure probabilities from 0.575 to 0.66. The results also suggested that during this landing, the pilot is not overloaded on the "downwind leg" of the approach.

Results, Launching Task

As in the landing task, the values of M, T, and F were found to have a pronounced effect on the number of subtasks which were completed during the simulation runs. In the worst cases encountered, the task was failed by the simulated operator after completion of as few as eight subtasks. In the best case, no launching failures were encountered for an entire run. A summary of the results concerning task completion is presented in graphical form as Fig. 3 for all runs; here the number of task failures for each run are plotted as a function of M. The effect of changes of T and F may also be determined from this figure. Outside criterion data for the launching task indicated that 16 out of 88 attempted launchings in the real situation resulted in failure. This figure is indicated in Fig. 3 as an arrow on the ordinate. Inspection of Fig. 3 indicates that the $F = 1.00$ curves do not intersect the outside criterion at any point. On the other hand, the $F = 0.90$ curve crosses reality and underpredicts the actual number of launching failures at M values between 1.95 and 2.80. An interpolated curve of $F = 0.91$ (dotted curve, Fig. 3) would coincide with the outside criterion data with M values in the specified range.

These results may be interpreted as indicating reasonable concordance between the results of the two applications of the model. As stated above, it is noted that data from Fig. 3 would in fact cross the outside criterion at a value for M, which lies within the range of values predicted from the landing task. Thus, agreement would accrue if the missile launching squadron was 4.5 per cent faster than average and the landing squadron was 4.5 per cent slower than average.

It is also known that the outside criterion data for the missile launching task were obtained during daylight missile test situations in which the attack was staged; the carrier launching data were obtained under more

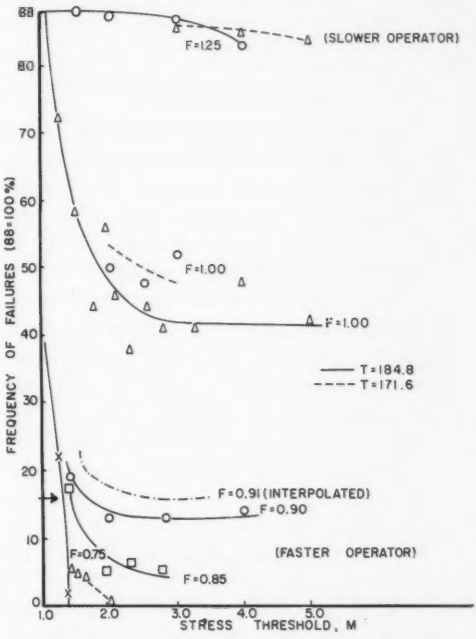


Fig. 3. Frequency of task failures, launching task.

varied and realistic operational conditions. The favorable missile test flight conditions may account for the comparatively low failure rates shown by the pilots involved in the launching analysis.

Comparison of Figs. 2 and 3 indicate that:

(1) the model yields relatively smooth curves whose form is characteristic and repeatable from one task to another;
(2) the data serve to confirm the gross expected results that higher success probabilities are obtained by faster operators (smaller F values), by operators who can

tolerate more stress without breaking (higher M values) or by an increase in the time limit (larger T value);

(3) for the average pilot in the launching task ($F = 1.0$) a failure probability between 0.40 and 0.435 was indicated in the stress threshold range coincident with that previously derived for landing aboard a carrier.

CONCLUSIONS

It may now be concluded that over two tasks the model (1) gave results which agreed in the large with information received from pilots; (2) was generally continuous and consistent; and (3) predicted reality reasonably well. It appears that the goal of prediction of the effectiveness of the human operator in man–machine systems has been achieved to some extent for single-operator situation. However, refinements in the model may still be useful and future investigations of this type will be required in which the collection and classification of outside criterion data will be more closely controlled.

It is not possible to say as a result of the two tasks analyzed that the results are sufficient to indicate the model may now be applied in practice with sure success. The present effort has given reasonable data and it is considered that the model may now be utilized effectively for comparative evaluation of alternative systems designs in terms of operator loading.

Acknowledgement

We are especially indebted to Dr. D. Smith, Dr. M. Lund, and Dr. P. Cheatham for their advice and encouragement throughout this program. We would also like to acknowledge the contributions of Mr. Albert Meyerhoff, who provided some of the basic concepts incorporated into the model during its formulation and to the staff of the Aeronautical Computation Laboratory of the U.S. Naval Air Development Center, especially Dr. Morris Plotkin, Mr. Joseph Porter, and Mrs. Mary Knight, who performed some of the mathematical analyses and all of the computer programming. The computer used to perform the calculations described in this report was the Model 650 IBM Magnetic Drum Data Processing Machine. Computations were performed both at the Aeronautical Computation Laboratory, U.S. Naval Air Development Center, Johnsville, Pa., and at the Steam Turbine Division of the Westinghouse Corporation, Lester, Pa.

REFERENCES

1. KAHN, H., *Applications of Monte Carlo*, O ce of Technical Services, Department of Commerce, Washington 25, D.C. Pp. 39–41, 238–241, 27 April, 1956.

2. STIBITZ and LARRIGEE, *Mathematics and computers*. McGraw-Hill, New York, 1957.

Operator Accuracy in Angle Measurement and Transfer

ROBERT C. HAYGOOD and RICHARD K. OVERTON, Autonetics, a Division of North American Aviation, Inc., Downey, California.

A series of three experiments was conducted to determine operator accuracy in the measurement and transfer of angles with a first-order theodolite. Accuracy was expressed in terms of the variability of operators' measurements and transfers. Experienced operators showed a standard deviation of measurements of 0.51 sec and a standard deviation of transfers of 2.89 sec. Newly trained operators showed an estimated standard deviation of measurements of 0.57 and 0.52 sec in two successive measuring sessions. Results with newly trained operators indicated that operators who wear glasses with satisfactory correction may be considered equivalent to operators with "normal" eyesight. Typical measurement and transfer procedures are discussed and recommendations for modification of these procedures are given.

INTRODUCTION

A necessary step in the activation of any inertial guidance system is the alignment of the system's stable platform. The accuracy of this alignment sets a limit on the accuracy with which the system will function during a mission, since the system will compute its course and position with reference to the original alignment.[1]

In general, the alignment process consists of some combination of two essential operations: the measurement of an unknown angle, and the transfer, or "setting in" of a known angle. These operations currently are performed with optical instruments by a human operator using techniques analogous to those employed in surveying. From a practical point of view, a rather precise evaluation of the error in these operations is necessary in predicting missile accuracy.

The experiments reported herein were designed to secure data on three points: (a) Operator accuracy in measuring and transferring angles, (b) Relative accuracy of experienced and newly trained operators in angle measurement, (c) Relationship between visual acuity and accuracy.

Because the concern of the experiments was with operator error rather than instrument error, the mean and variability of the operators' measurements were chosen as the criteria of accuracy. With such criteria, the effects of a constant instrument error are cancelled out, and the question of the "true" size of the angle, which otherwise would be quite crucial to the experiment, is eliminated. A thorough discussion of instrument errors has been given by Whitehead.[4]

GENERAL PROCEDURE AND APPARATUS

Optical instruments. The instrument typically used for precise angle measurement is a theodolite. Basically, it consists of a telescope which is linked optically to a protractor-like scale. When the telescope is centered on a target, the direction of that target can be read from the scales. The difference of the scale readings of two targets thus forms a single *observation* of the size of the angle between them. In this report, the mean of a series of such observations is called the operator's *measurement* of the angle.

The telescope can also be turned from a reference target, of known direction, until the scales read a desired value; a target can then be placed so that it is centered in the telescope. This makes it possible to *transfer* an angle, that is, to position or align other equipment with the theodolite.

Two different commercially available theodolites were used in the present study. Both were "first-order" instruments of the type used when the greatest possible accuracy is required. Under the conditions of these experiments, no significant differences were found between instruments. Therefore, the data were grouped without regard to instrument used.

EDUCATION LIBRARY

Gauss eyepiece. Each theodolite was equipped with a Gauss eyepiece. This type of eyepiece projects a pattern of light from the telescope; the pattern can be reflected back into the telescope from a mirror. When the reflected pattern is centered in the telescope, the telescope line of sight is perpendicular to the mirror.

In measuring an angle, the telescope is made perpendicular to the mirror. Therefore, the direction of the telescope is determined by the orientation of the mirror, and the angle between the perpendicular to the mirror and the line to another target can be measured.

In transferring an angle, the telescope is first turned until the scales read a desired value, then the mirror is made perpendicular to the telescope line of sight. Thus a piece of equipment with a mirror mounted on it can be aligned in any chosen direction.

Targets. In addition to mirrors, the reticles within another theodolite can be used as a target. The lenses of the theodolite collimate, or make parallel, the rays of light leaving the telescope. Therefore, both the reticle target and the reflection from the mirror require infinity focus. This means that both types of targets can be placed as close as desired to the operator's theodolite without creating any focus problem, and that the focus need not be changed while measuring or transferring an angle.

Measurement procedure. The experiments reported herein were concerned only with horizontal, or azimuth, angles. The reason for this is that long-range missiles are customarily launched vertically, and sufficiently accurate electro-mechanical means exist for aligning guidance systems to the local vertical.

The multiple reiteration method of angle measurement was used in all experiments.[2] In this method, many observations are made. Half of these are made with the theodolite in its "direct" position and half in its "reversed" position. To go from direct to reversed position, the operator performs a maneuver known as "plunging and reversing", which effectively turns the telescope upside-down and reverses right and left. The effect of matching each direct observation with a reversed observation is to cancel out errors due to displacement of the telescope reticles, eccentricity of the theodolite scales, and the like. In addition, the horizontal scale is rotated within the theodolite before each observation. This insures that the observations are made over different portions of the scale, and that errors due to imperfect scale markings are largely cancelled out.

The operator first places the telescope in direct position and centers the telescope on the left target. Then he sets the horizontal scale approximately to a reference value. Next he adjusts the vernier dial and makes the left target scale reading. Turning the telescope to the right target, he leaves the scale in the same orientation, adjusts the vernier, and makes the right target scale reading. The difference between the left and right target scale readings forms his first observation of the angle in direct position.

Then the operator plunges and reverses the telescope, centers on the right target and sets the horizontal scale to the same reference value used for the previous direct observation. After adjusting the vernier and making the right target scale reading, he turns to the left target and makes the left target scale reading. The difference of the scale readings for the two targets forms his first observation in reversed position. The first pair of observations is then repeated as many times as desired, with a new reference value for each pair.

If there is a defect in the instrument, observations made with the theodolite in its direct position may be consistently larger or smaller than observations made in its reversed position. Such a defect would, of course, make an individual's observations more variable than they otherwise would be. To prevent the appearance of spuriously high variances in these experiments, each direct observation was paired with its corresponding reversed observation. Then the pair was averaged. Such averages were used in computing the variance within a set of observations. All measurement results

in this report were derived from observations paired in this manner.

Transfer procedure. In contrast to the reiteration method used in measurement, transferring is a "one-shot" operation. Each transfer destroys the alignment established by any preceding transfer, and the benefit of averaging cannot be obtained.

The operator first centers the telescope on a reference target and sets the horizontal scale approximately to a convenient value, such as 180° or 0°. Then he adds the size of angle to be transferred; thus he finds the required final scale reading. Next he turns the telescope until the required final scale reading shows in the scales. The final step in the transfer is to adjust a target mirror until the return Gauss image is centered in the telescope.

EXPERIMENT 1

This experiment was designed to determine the standard deviation of experienced theodolite operators' measurements of an angle.

Method

Subjects. Fifteen experienced theodolite operators served as subjects in this experiment. Their experience ranged from three months to several years. All subjects were employees of Autonetics and were excused from normal work duties to participate.

Equipment. The right target in Experiment 1 consisted of the reticles within a theodolite located about 14 ft from the operator's theodolite. The left target was a 4.5-in. optically flat mirror 12 ft from the operator's theodolite. Both targets and the measuring theodolite were mounted on heavy metal stands. The experiment was conducted in an air-conditioned laboratory with a reinforced concrete floor. The experimenter prepared the operator's theodolite for use prior to the arrival of the subjects.

Procedure. When each subject reported for the experiment, the experimenter explained the purpose of the experiment, and told the subject that the angle between the target theodolite and the target mirror was to be measured. The instructions stressed accuracy rather than speed. The multiple reiteration method of angle measurement was used, with the subject operating the theodolite and the experimenter serving as recorder, that is, writing down the subject's readings. Sixteen pairs of observations were made by each subject, each pair starting with a new reference value. All computations were performed by the experimenter following the experiment.

Results and Discussion

Gross errors. When an operator misreads the scales, or bumps the instrument during an observation, a value far outside the normal range of variability may be obtained. In this experiment, if a particular observation differed by more than 10 sec from the mode of all subjects' observations, it was arbitrarily said to be a "gross error". The gross errors constituted 11.8 per cent of the observations made in the experiment, and ranged up to 100° in size. Gross errors were excluded from all further analyses.

Standard deviation of measurements. As noted above, the primary goal of this experiment was to determine the standard deviation of measurements made by a large number of experienced theodolite operators. Since only fifteen subjects participated, appropriate statistical procedures were used to estimate the results that would have been obtained if a large population of operators had been tested. The obtained sample results and the estimated population values are presented in Table I.

TABLE I

Results of Experiment 1

Accuracy	*Sample value*	*Estimated population value*
Variance of measurements	0.26 sec^2	0.28 sec^2
Standard deviation of measurements	0.51 sec	0.53 sec
Mean standard error (estimated standard deviation of measurements)	0.51 sec	—

Serial effects. To evaluate the cumulative effects of fatigue and other factors, each of the sixteen pairs of observations was considered as a trial. The variance across subjects on each trial increased somewhat irregularly as the trials progressed, reaching a maximum of Trials 11, 12, and 13. This suggested that perhaps very little additional accuracy was added by the last eight trials.

To evaluate this hypothesis, the standard errors of trial means for Trials 1–8 and 1–16 were compared. If the longer series of trials was, in fact, a more accurate measurement, its standard error should have been smaller. However, the standard error for Trials 1–8 was 0.11 sec while the standard error for Trials 1–16 was 0.13 sec. Thus it appears that the mean at the end of sixteen trials may well be less accurate than that at the end of eight. For this reason, it was recommended that eight pairs of observations be made instead of sixteen.

In practical alignment situations, time is often critical. Cutting angle measurement time roughly in half without a significant loss in accuracy therefore would represent a substantial gain.

Memory effects. An operator can, through fairly simple mental arithmetic, derive some idea of the approximate range of his early observations, and restrict later observations to fall well within his estimate of this range. The distribution of ranges for the first six and the last six trials was compared by the Kolmogorov–Smirnov method (one-tailed). No significant restriction of range was found ($KR = 0.43$, $p > 0.10$), and the memory effect hypothesis was not supported.

Speed and accuracy. Mean time taken by subjects was 96 min, exclusive of instruction time at the beginning. Subjects differed widely in speed, as evidenced by a standard deviation of 25 min.

When accuracy was defined as the difference between the subject's measurement and the group mean, the correlation between speed and accuracy was not significant ($r = +0.10$, $p > 0.10$). It should be noted again, however, that instructions stressed accuracy rather than speed.

EXPERIMENT 2

The purpose of this experiment was to determine the standard deviation of experienced theodolite operators' transfers of an angle.

Method

Subjects. Seventeen experienced theodolite operators served as subjects. As in Experiment 1, the subjects were all Autonetics employees and were excused from normal work duties to participate.

Equipment. The experimental conditions were substantially the same as in Experiment 1, except that the target mirror was set up 2 ft from the operator's theodolite to facilitate adjusting the mirror while looking through the telescope.

Procedure. When each subject reported for the experiment, the experimenter explained the purpose of the experiment, and asked the subject to transfer an angle of 55° 17′ 6.0″ from the reticle target on the right to the target mirror on the left. The transfer method outlined above was used, with the experimenter assisting the subject in the calculations. Instructions again stressed accuracy rather than speed.

In addition, the second through eighth subjects were also asked to measure the transfer left by the preceding subject. Because of scheduling difficulties, and the undesirability of allowing the transfer to stand for a period of time before measuring, the experimenter performed the measurement of the transfers made by the eighth through seventeenth subjects. No significant difference was found between the variability or mean of measurements made by the subjects and the experimenter. The multiple reiteration method was used for measurements, with eight pairs of observations being made. The change to only eight pairs of observations was suggested by the results of Experiment 1.

Results and Discussion

Gross errors. Two of the subjects made gross errors in transfer; that is, the measured values of the angles they transferred differed by more than 10 sec from the desired 55° 17′ 6.0″. The results of these subjects were excluded from all other analyses.

The percentage of gross errors in transfer was 11.8. Gross errors also appeared in 11.8 per cent of the observations of the first experiment. In that experiment, however, they were scattered through the observations of almost all the subjects. It is probably only a coincidence that the two percentage figures are identical.

Standard deviation of transfers. Since each transfer was measured to determine the subject's success, the total variance (s_T^2) included the variance of measurements (s_m^2) and the variance of the transfers (s_a^2). Because of the nature of the task, the accuracy of a measurement or transfer is independent of the size of the angle, and

$$s_T^2 - s_m^2 = s_a^2$$

However, s_m^2 is unknown, and must be estimated before s_a^2 can be found. This can be done either (a) by using the s_m^2 from Experiment 1, or (b) by estimating from the variance within each measuring operator's set of observations. Since the standard error computed for each measuring operator is an estimate of how far his mean is likely to be from the "true" mean (in the absence of personal bias), the mean standard error ($\overline{SE}_w$) for measuring operators should approximate the standard deviation of the operators' measurements, or

$$\overline{SE}_w \cong s_m$$

After analysis of the task, it was concluded that personal biases were unlikely, and that such biases as existed would cancel out over a large number of operators. This conclusion was subjected to empirical test and confirmed in a striking way in Experiment 1; both $\overline{SE}_w$ and s_m were 0.511 sec. It should be noted that 0.001 sec is 1/129,600,000 of a circle.

The total variance for the fifteen subjects who did not make gross errors was 8.74 sec²; $\overline{SE}_w^2$ was 0.29 sec². The resulting s_a^2, or variance of transfers, was thus 8.35 sec² with a corresponding standard deviation of transfers of 2.89 sec. Using the measurement data of Experiment 1 instead, the standard deviation of transfers was 2.91 sec. The difference of the two estimates was both practically and statistically insignificant.

Speed and accuracy. The time required to complete the transfer ranged from 6 to 12 min, with a mean of approximately 8 min. As in Experiment 1, no relationship was found between speed and accuracy.

EXPERIMENT 3

This experiment was designed to determine the standard deviation of newly trained theodolite operators' measurements of an angle, and the relationship of visual acuity and accuracy.

Method

Subjects. Thirty engineers and other technical staff members of the Autonetics System Test Department served as subjects. None had previously operated a theodolite. The participation of these subjects was scheduled by the Department, and was considered a part of the training of the personnel involved.

Equipment. The experimental conditions were substantially the same as those of Experiment 1, except that the targets shifted several times. These shifts were presumably due to intermittent heavy ground traffic in the area adjacent to the experimental laboratory. They continued despite several attempts to secure stability by rearranging the equipment.

Familiarization. The subjects were divided into groups of seven for purposes of orientation. These groups were given a one-hour lecture on the principles of theodolite operation, and a demonstration of the instrument. Two subjects did not receive the group lecture, and were instructed individually. Following the orientation meeting, each subject was scheduled for two individual sessions during the succeeding four or five days.

Procedure. During the first individual session,

the experimenter continued the familiarization, pointed out the location of the various theodolite controls, and allowed the subject to operate these controls. When it appeared that the subject was sufficiently familiar with the instrument, he was asked to measure the angle. During the second session, the subject was allowed to begin angle measurement immediately. The multiple reiteration method of angle measurement was used, and each subject made eight pairs of observations during each session. The experimenter kept a close watch on each subject and frequently interrupted to give further instructions and to clear up procedural points.

Results and Discussion

Gross errors. In this experiment, a gross error was defined to be any observation which differed from the subject's own median by more than 10 sec. The experimenter performed most of the calculations while recording the subject's readings. When a gross error was detected, the subject was required to repeat the observation, and a notation was made in the record. During the first session, 6.9 per cent gross errors were made. During the second session, 4.8 per cent gross errors were made.

Estimated standard deviation of measurements. The shifting of targets precluded the possibility of computing the standard deviation of measurements directly. However, no significant shifts occurred during any one subject's set of observations. Therefore, the method of estimation used in Experiment 2 was followed; that is, the standard deviation of measurements was estimated from the variance within each subject's set of observations. As previously noted, considerable confidence is placed in these estimates for the particular procedures employed in these experiments.

Group performance, as estimated from the average of individual performances, is summarized in Table 2. The results of all experiments are included for comparison. None of the differences shown was significant.

A somewhat surprising result of this experiment was the finding that a group of subjects with no previous experience could perform as well as experienced operators after only a brief training period. The total training time for each subject prior to his first angle measurement was about 1½ hours.

Before generalizing to practical situations, however, several points should be noted:

1. All subjects were graduate engineers or

TABLE II

Accuracy and Speed Data from all Experiments

Angle measurement	*Experiment 1*	*Experiment 2*	*Experiment 3*	
			Session I	*Session II*
Estimated standard deviation of sample	0.51 sec	0.62 sec	0.57 sec	0.52 sec
Mean standard deviation within each subject's set of observations	1.99 sec	1.64 sec	1.49 sec	1.40 sec
Number of pairs of observations	16	8	8	8
Number of subjects	15	17*	30	30
Mean time to complete one pair of observations	6.0 min	not recorded	9.0 min	7.1 min
Angle transfer				
Standard deviation of transfers	—	2.89 sec	—	—
Mean time to complete transfer	—	8.0 min	—	—

* 17 subjects transferred angles. Measurement data are based on 7 of these who also measured angles. The experimenter measured the remainder of the angles.

experienced technical personnel. They therefore are undoubtedly of somewhat higher intellectual levels than the typical randomly selected group. Whether this difference in intellectual levels might affect the variables in question can only be determined by further experimentation using subjects drawn from the general population.

2. An understanding of theodolite operation is essential to the activities in which these personnel will engage. The subjects therefore showed unusually high interest in learning about theodolites and in performing well.

3. Subjects were not asked to learn how to prepare the instrument and targets for use or how to perform the necessary computations. Sufficient training to prepare the subjects to perform these tasks without supervision would undoubtedly require much more time. Angle measurement itself, however, is apparently a relatively simple task which could be performed by an intelligent person under supervision after relatively little training.

Visual acuity and accuracy. To evaluate the hypothesis that accuracy is dependent on visual acuity, the experimenter examined the results of the subjects' eye examinations taken at the time of employment. These records showed that every subject had either "normal" eyesight or a satisfactory correction to "normal". Consequently there was no relationship between corrected eyesight ratings and performance in the experiment. Records of uncorrected visual acuity were incomplete, but those available showed no relationship to performance. Performance of operators who wore glasses was the same as that of those who did not. These findings indicate that operators whose glasses provide satisfactory correction can probably be considered equivalent to operators with "normal" eyesight.

OPERATIONAL PROCEDURE RECOMMENDATIONS

The following recommendations for modification of the theodolite operating procedure were generated as a result of this series of experiments:

1. The recommendation for reducing the number of pairs of observations from sixteen to eight was mentioned above. If increased alignment error can be tolerated, and alignment time is critical, consideration should be given to reducing further the number of pairs of observations to four. On the basis of figures available, it appears that this would increase the standard deviation of measurements from 0.5 sec to about 1.0 sec.

2. The regular multiple reiteration procedure requires that the operator plunge and reverse the theodolite after each individual observation. Although each direct observation should be matched with a reversed observation covering the same portion of the horizontal scale, it is not necessary to make the two observations in sequence, except for technical reasons in stellar observations.[3] Therefore, it is recommended that the operator make all direct observations first, then, after plunging and reversing only once, make all the reversed observations. This will reduce both the complexity of the task and the time requirement.

3. The procedure of having the recorder perform the computations as he records the subjects' readings has worked out quite well. The experimenter was able to catch all but six gross errors as they occurred in Experiment 3; with 756 individual observations this represents a total of only 0.8 per cent gross errors remaining in the data. The remaining gross errors were corrected immediately. It is recommended that this procedure be instituted wherever possible.

REFERENCES

1. O'Donnell, C. F., Inertial navigation. *J. Franklin Inst.*, **266,** 357–277, 373–402, 1958.
2. U.S. Coast and Geodetic Survey. Special Publication 247, *Manual of geodetic triangulation.* USGPO, Washington, 1959.
3. U.S. Coast and Geodetic Survey. Special Publication 237, *Manual of geodetic astronomy.* USGPO, Washington, 1952.
4. Whitehead, T. N., *The design and use of instruments and accurate mechanism.* Dover, New York, 1954.

Anthropometric Photogrammetry as Applied to Escape Capsule Design

JOHN W. CHAFFEE, Human Factors Laboratory, Convair—A Division of General Dynamics Corporation, Ft. Worth, Texas.

An experimental investigation was made of the location and spatial requirements of salient anatomical features of the human operator of a high-performance air weapon system when simulating the use of an escape capsule type of emergency abandonment system.

Twenty-four subjects representative of the Air Force flying population were measured using a new system of anthropometric assessment which employs a non-stereographic photogrammetric treatment requiring two to three ordinary 8 in. × 10 in. view cameras together with adequate stroboscopic illumination.

Data on the x, y, and z co-ordinates of sixteen anatomical features of the operators when positioned in six body attitudes representative of the escape system's use are presented in the form of figures.

"The design of an item whose success is dependent in any degree on the human effort expended on it must conform, configuration-wise, to the physical limitations of the human body."

The safety of the human occupant of high-performance aircraft has become a problem of steadily increasing importance. Many features to be found in modern aircraft cockpits are provided solely to protect and aid the occupant during systems malfunction and subsequent emergencies either on the ground or when airborne.

The most recent of these safety features is the escape capsule. Typically, the escape capsule is a device for completely enclosing the human operator prior to and during assisted abandonment (ejection) of a high-speed military aircraft. It provides a container for packaging the man in an environment which will briefly sustain biological needs against adverse conditions of low atmospheric pressures and temperatures, strong wind-blast, and hostile ground environmental conditions.

The design of such a device is quite complex, requiring considerations of many variables among which size and habitability are of especial interest to military and industrial physical anthropology.

Ultimately the size of an encapsulating device designed to accommodate humans is dictated, at least with respect to minimum internal space, by the range of sizes of the human bodies to be contained therein. These spatial requirements are, in turn, a function not only of human body size variability *per se*, but also of the particular attitude in which the body must be positioned for ejection.

Once the body attitude to be adopted by the occupant of the enclosed capsule has been established, it then becomes the task of the physical anthropologist to apply if possible existing data and/or techniques of anthropometric investigation to determine which body measurements are adequately descriptive and from what population they must be sampled. These anthropometric data will then be used to establish the minimum internal dimensions of the capsule, as well as the most desirable location for arm rests, head rests, and such restraint devices as are required. These data will also specify the location of the individual's eyes (for orientation of displays), principal joint centers (for kinematical weight and balance studies), and certain workspace areas (to assess amount of possible controls movement).

At the beginning of the present study physiological and engineering studies had established in general the limits of variability in body posture during emergency ejection as well as the basic seat-floor geometry.

It was assumed that the capsule was to be used by the particular segment of the military population described in the 1950 anthropometric survey of the "Air Force Flying Population".[10] Consequently it was decided that only

those individuals should be measured who were within the range of body size described in that survey.

The problem of which measurements to take was solved by a preliminary analysis of the bodily motions of several individuals while they went through mock procedures of actuating the encapsulating controls and ejection controls.

A highly generalized abstraction of possible future hardware was employed for this and may be seen in schematic in Fig. 4.

Of particular interest in this preliminary investigation was identification of the various different bodily movements required, for instance, in reaching for and elevating simulated actuation control handles and, at the same time, retracting the legs to the rear and positioning the head and torso in an optimum manner.

Examination of photographs taken of these programmed movements revealed that the range of significant changes in bodily configuration could be summarized by three discrete postures shown in Fig. 1 (A), (B) and (C).

In Fig. 1 (A) is illustrated the preliminary attitude typical of the beginning of the movement sequence in which the operator reaches for the actuating controls which will close the capsule shell.

At this time the head is placed firmly against a surface which is contiguous with the seat back. The feet are either still on the rudder pedals (in the case of the pilot or co-pilot) or at some nominal point on the floor.

Fig. 1 (B), the ejection attitude, shows the legs fully retracted. The encapsulating controls have been fully actuated by being elevated and swung to the rear. In the present instance, these controls were attached to a hinged section of the seat bottom and thereby raised this surface when actuated. This is the body attitude typical of actual emergency jettison or ejection.

Fig. 1 (C) depicts the body attitude typical of encapsulated flight. Here it is assumed that encapsulation has occurred but actual ejection may not necessarily be required.

In the case of the preliminary attitude [Fig. 1 (A)], we are interested in discovering the best position and angle for installing the encapsulating controls and in determining necessary hand clearances.

On the other hand, in the case of the ejection attitude shown in Fig. 1 (B) many problems required answers. For example, the most desirable locations for a head rest and for arm rests was of concern. Further, the location and necessary lateral, vertical, and fore-aft clearances for feet, knees, knuckles, elbows, shoulders, and head were extremely important.

Finally, in the encapsulated flight attitude of Fig. 1 (C), in addition to the requirements in common with the ejection attitude for locating foot, knee, shoulder, and head clearances, there existed the additional need to stipulate the mean locus of the operator's eye in order to locate the visual field. Also, some estimate of the amount of lateral movement possible with the flight control stick was necessary.

During the course of the preliminary investigation it was discovered that variation in the height of the Seat Reference Point (SRP) above floor or Heel Rest Line (HRL) produced considerable changes in foot and knee clearances. This can be appreciated by inspection of Fig. 2 where the effect of three SRP-HRL distances upon vertical and horizontal clearances of the knee and toe is qualitatively shown. This figure is a composite tracing of three photographs of the same individual in the ejection attitude.

In each case, the subject shown in Fig. 2 is attempting to force his legs and feet as far to the rear (aft) as possible.

Consideration of SRP-HRL distance as a variable and not a fixed value was thus required as it represents the adjustability normally available in aircraft seats[5] and constitutes a slightly complicating circumstance in the analysis of body clearance data such as were collected.

It was decided therefore to determine the effects of variability in SRP-HRL *per se* upon body measurements by selecting four discrete

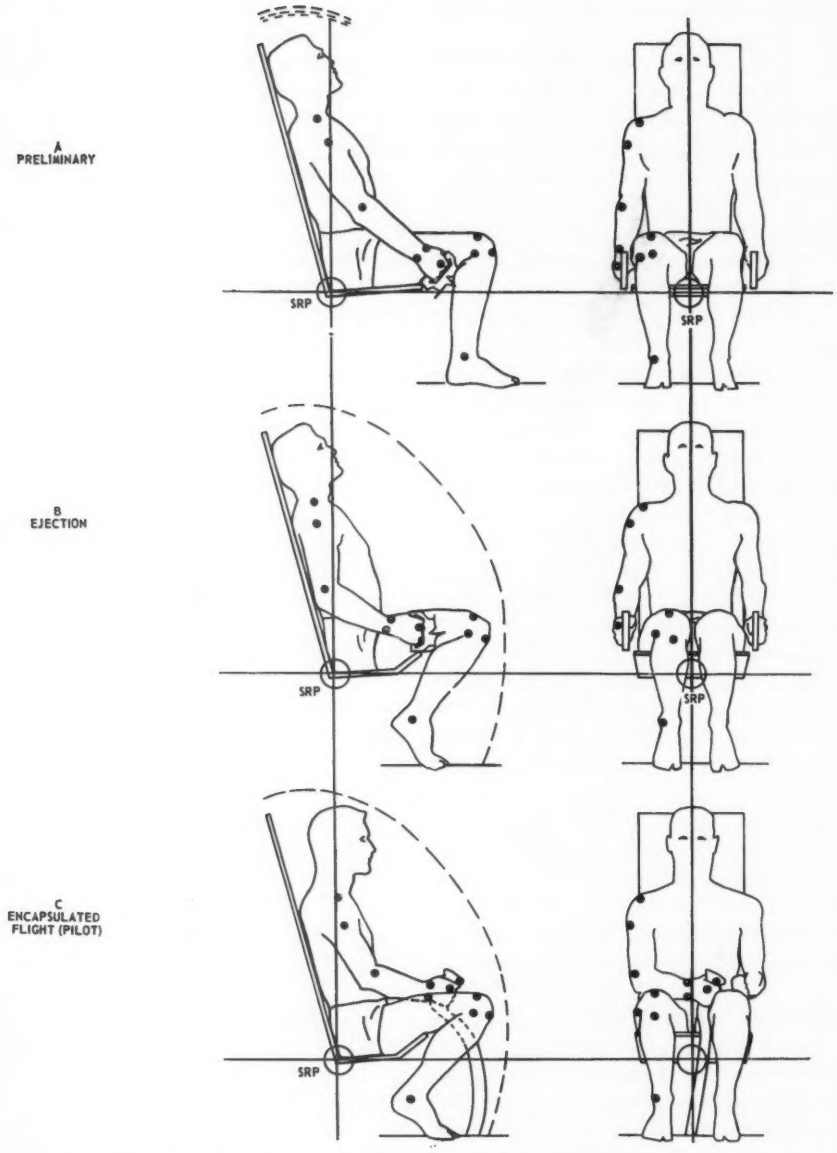


Fig. 1. The three critical body postures which are representative of the encapsulating systems sequence of body positional changes.

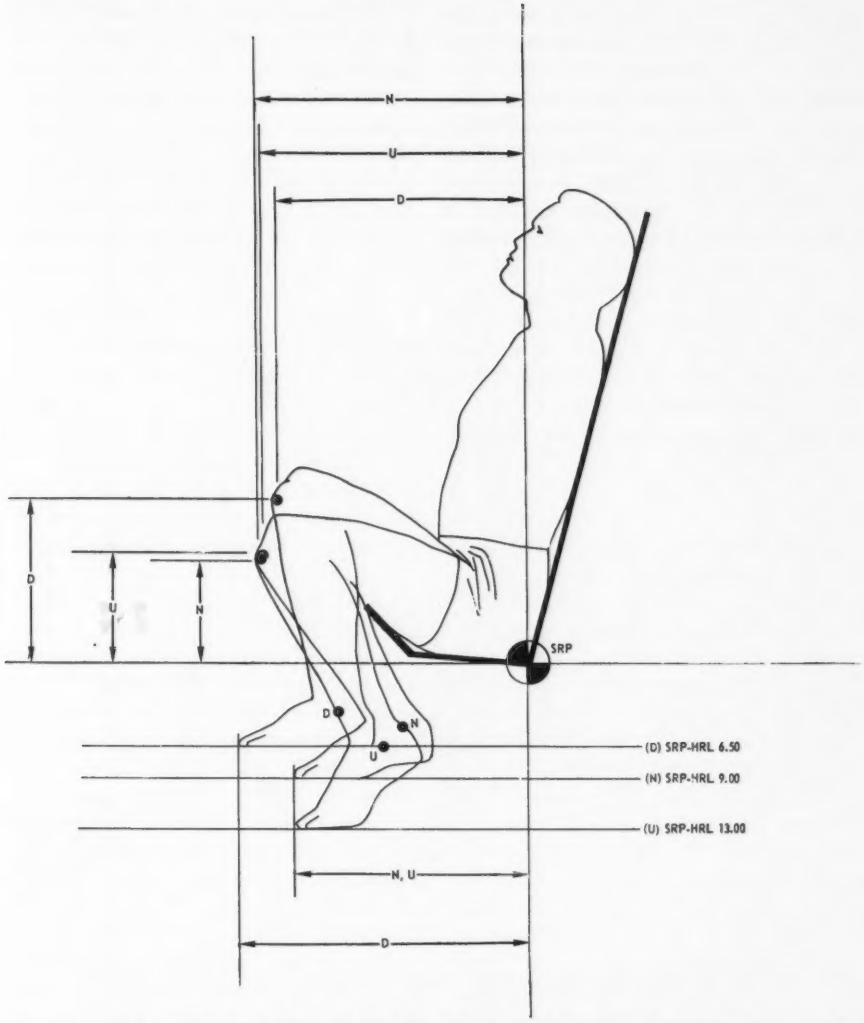


Fig. 2. A composite tracing of three photographs taken of progressive increase in SRP-HRL distance. Changes in loci of markers on knee and ankle as a function of SRP-HRL are indicated.

levels of that variable and observing each subject at each level.

Presently available civilian[12] and military anthropometric data[5, 8, 10, 16, 17] are not descriptive of pilot's body attitudes adopted in the aircraft nor particularly those of interest in the present study.[4]

Indeed, it is difficult to find in the literature of military anthropometry even a rationale for the acquisition of the type of data needed in this study.

Davenport has commented that ". . . we have (for example) a satisfactory method of measuring nose height but no satisfactory method of indicating the position of the nose on the head. . . . And so with practically all

measures (apart from some of those in the sagittal plane), their relation to absolute planes of reference is not determined."

"For to determine completely the position of any point, it is necessary to measure its co-ordinates from each of three perpendicular planes. Nothing less will suffice; more is superfluous. This is the co-ordinate principle of Déscartes' which has been successfully utilized for over a century and is the basis of all engineering projects."[7]

In view of the lack of pertinent data or a suitable measurement rationale and operational technique it was determined that an original data acquisition approach was needed—one which produced data which are:

1. Based upon three-dimensional space.
2. Verifiable.
3. Accurate.
4. Precise.

To this end, a facility similar to that shown in Fig. 3 was developed whereby an accurate photogrammetric record of body size and configuration could be obtained. For this, commercially available 8 × 10 Burke and James View Cameras fitted with long focus lenses (Goertz "Red Dot" Artar, F.11, 19 in. focal length) were used in conjunction with eight 100 watt/sec electronic stroboscopic flash units. Stroboscopic lights synchronized with the camera shutters were preferred since they "freeze" the subjects' movements such as breathing and muscular twitch. Further it is only with stroboscopic lighting that the high light levels can be comfortably withstood and cheaply obtained which are required with the small apertures (F.90) used.

It should be pointed out here that the arrangement used was instrumented with relatively cheap off-the-shelf commercial equip-

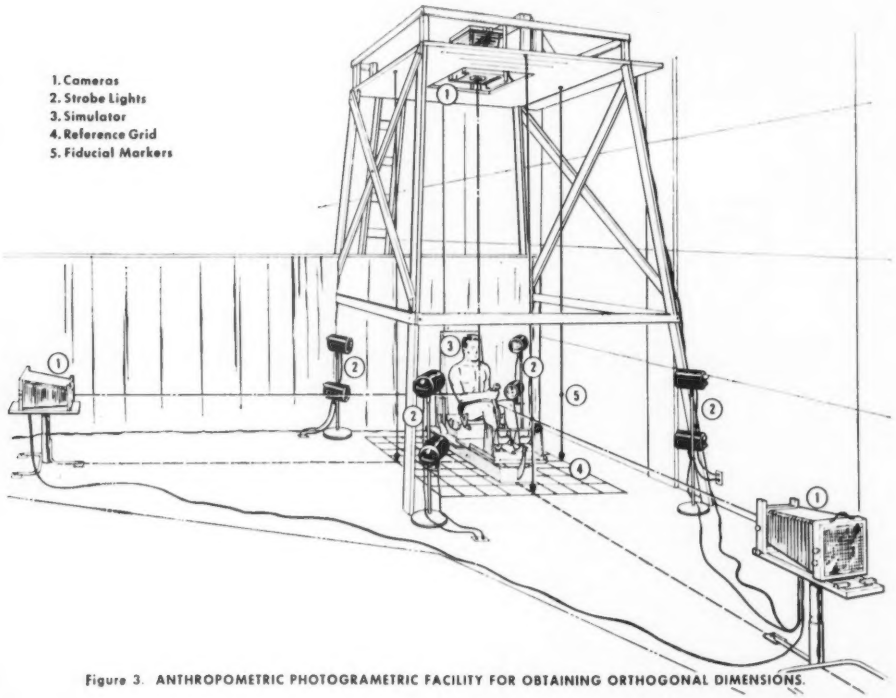


Figure 3. ANTHROPOMETRIC PHOTOGRAMETRIC FACILITY FOR OBTAINING ORTHOGONAL DIMENSIONS.

ment. No specialized equipment or procedures often associated with previously reported stereophotogrammetric studies[11,14] are required since stereoscopic viewing is not involved in this technique.

Our method produces three 8 × 10 negatives on which appear the test subject, photogrammetrically accurate fiducial marks and Seat Reference Point (SRP). These negatives are measured by hand with either commercially available precision steel scales graduated to 0.01-in. divisions or as in the present study with a semi-automatic device such as the Benson–Lehner X–Y Oscillograph Reader ("Oscar").

The position of each salient anatomical feature previously marked on an individual appears on the negatives and is measured in analog space with respect to the apparent horizontal, vertical, and lateral displacement from the SRP image.

These data describing the "apparent" displacements are then operated upon by appropriate transformation functions to produce true x, y and z Cartesian co-ordinates of the anatomical feature's locus in space relative to the SRP. Thus a complete co-ordinate dimensional description is possible for each point.

Since the negatives form a permanent record, subsequent analysis of them may be performed to obtain various type data other than that originally required. Also check may be made in cases of apparent error in tabulation—a common occurrence in survey-type work.

In conjunction with this photogrammetric facility, an adjustable wooden mock-up was employed similar to that shown in Fig. 4. The purpose of this was, as mentioned previously to provide an abstraction of particular body support configurations.

Fig. 4 shows the principal features and dimensions of this device. These correspond in general with those recommended by the Air Force for "Stick-Controlled Airplanes".[1]

An exception to this standard may be seen as the hinged forward edge of the seat bottom. This is the forward 5 in. of the surface, which elevates from an initial angle of 5° to the horizontal through an arc of 42° to attain a final angle of 47° to the horizontal.

Elevation of this seat panel is occasioned when the actuating handles are forcibly raised by the operator and is one method of providing increased body restraint during ejection.

It will be observed that three SRP-HRL distances are shown in phantom in addition to the 13.00-in. level shown in solid outline. The manner in which the SRP height above the Heel Rest Line was varied might be of interest since it was necessary to maintain the SRP itself at a fixed point in space relative to the cameras.

The variability of SRP-HRL from 6.50 to 13.00 in. was effected by changing the elevation of the HRL, the movable platform, which is shown in Fig. 4. Upon this was mounted, as on the floor of an aircraft, a simulated flight control or "joy stick". No rudder pedals were necessary for this investigation.

The grip[13] on the flight control stick was the Air Force standard type B-9. The initials "GRP" represent Grip Reference Point which is the locus of the second finger (Digit III) when the hand is normally grasping the grip. The stick and grip were so installed as to locate the GRP 19.00 in. forward of the SRP and 22.50 in. above HRL when the SRP-HRL was 9.00 in.[3]

Although the SRP-HRL variable is in an aircraft continuous throughout a typical range from 6.50 through 13.00 in.[3] it was tested here at the discrete intervals of 6.50, 9.00, 11.50, and 13.00 in. The recommended range for a pilot's station is 6.50 through 11.50.[3] Some crew station seats go even higher, however, and so the range was extended arbitrarily to include a SRP-HRL of 13.00 in.

In Fig. 4 "grip area" represents schematically the locus and relative size of the area within which it was felt desirable to limit the testing of "optimum" location for the encapsulating controls.

There were no arm rests or leg restraints. No head rest other than the vertical extension of the seat back was used. No cushioning of

any type was employed since the bare wooden surfaces were supposed to simulate the top of an Air Force survival kit under a compressed or "bottomed out" seat cushion as well as the virtually non-compressible surface of standard back-type parachute.

Since it was of interest to learn the relative standing of the test subjects, with respect to "standard" measurements presently available on the Air Force flying population,[10] each subject was measured nude on a Morant type Grid[15] using landmarks and, as near as possible, methods described in the 1950 Air Force survey.

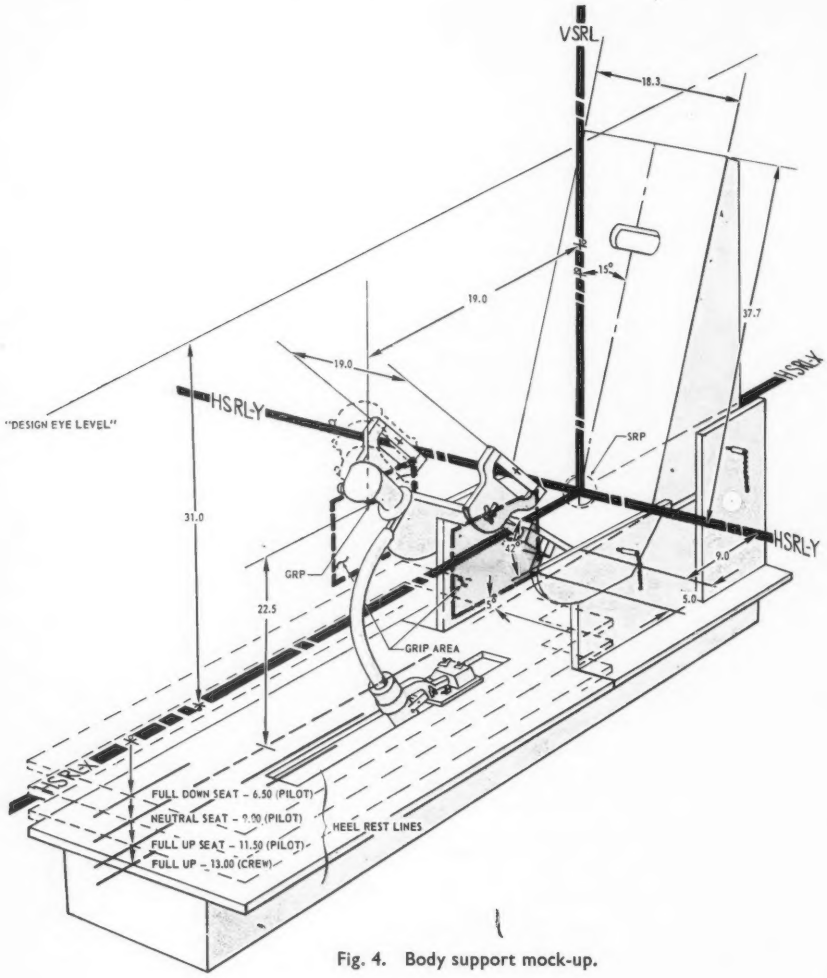


Fig. 4. Body support mock-up.

After this preliminary anthropometric assessment, the photogrammetric records were taken on the unclothed subjects in the following manner:

The wooden simulator was adjusted and the

preliminary body attitude of Fig. 1 (A) was simulated. The subject was suitably positioned in the mock-up and requested to grasp the loosely adjusted actuating control grips in a firm manner with arms and shoulders maximally extended.

immobility. Only the right-hand position and angulation were photogrammetrically recorded.

As suggested in Gavan's study,[9] notepaper reinforcing rings were used as markers. These were placed on the body surface at the points indicated in Fig. 5.

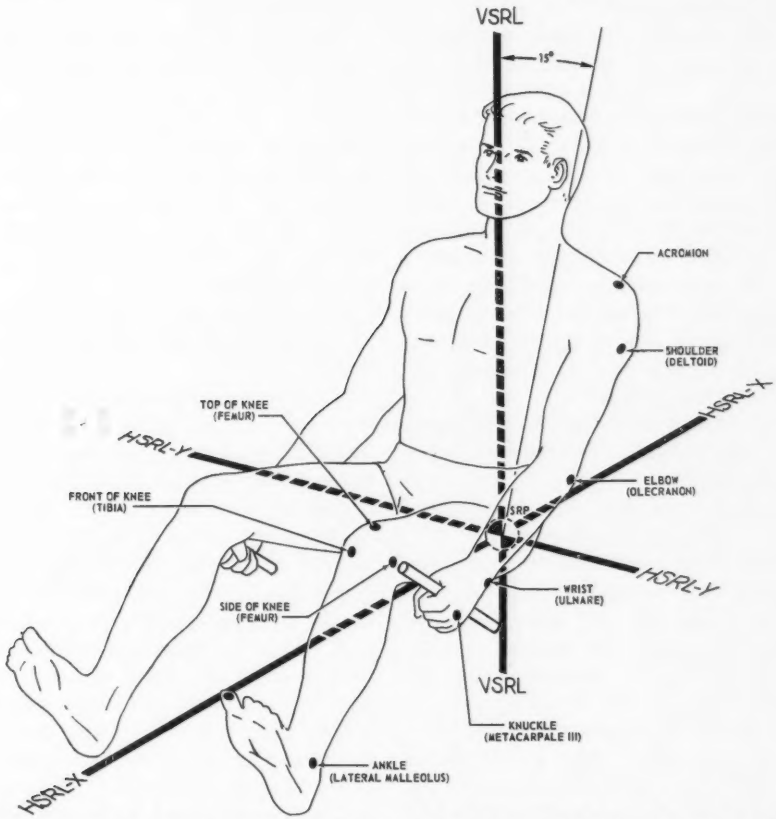


Fig. 5. Location of the pasted-on markers used to locate salient anatomical features, which are difficult to otherwise pinpoint in a photograph.

The subject was required to extend his arms and shoulders maximally as well as to deliberately vary the angle of the hand at the wrist in a random fashion until the preferred ("comfortable") angles for the two grips were attained.

At this point in the protocol the grips were locked by means of thumb-screws into rigid

When all was in readiness the signal was given to record the data, whereupon the synchronized cameras were tripped, firing simultaneously the stroboscopic lights. In this manner the necessary data on the subject's body size, locus, and configuration became a permanent record in a matter of $\frac{1}{3000}$ of a second.

The SRP-HRL adjustment for the preliminary attitude was 13.00 in. for all subjects. This was strictly for convenience since it had been determined previously that the SRP-HRL variable had no effect on body measurements other than those pertaining to the feet and legs. The data of interest here, however, did not concern the feet and legs.

While the cameras were being reloaded the subject was positioned in the ejection attitude of Fig. 1 (B). The subject was instructed to pull his feet as far back as possible while at the same time pulling up and back on the actuating controls. These controls, although now locked into a fixed positional geometry, rotated via the forward edge of the seat through a 42° angle to the elevated attitude shown in Fig. 1 (B).

The seat edge was then locked into the elevated attitude. The subject remained grasping the grips and exerting a small amount of force upwards and aft to assure realistic shoulder and arm geometry.

During the foot retraction and grip raising sequence, the subject's head remained firmly against the seat-back surface.

The cameras and light were then tripped obtaining thereby the required data.

While the cameras were again being reloaded, the subject was requested to relax his grip, move his arms, legs, and head about and generally relax. When it appeared that the subject was settled into a posture estimated as reasonably typical of that observed in pilots during extended flight, he was instructed to assume the encapsulated flight attitude shown in Fig. 1 (C). In this attitude, the subject was instructed to gaze straight ahead into one of the cameras. The right hand gripped the flight control stick which was fitted, as mentioned previously, with the Air Force standard B-9 stick grip. The stick was positioned in approximately the "neutral" position. "Approximately" is used here since the subject was actually instructed to shove the stick as far to the left as possible while maintaining a relatively midway position with respect to fore and aft "standard" stick travel. This stick position would produce an extreme left aileron input into a flight control system.

The left hand was placed either on the thigh or in the lap.

The floor was then raised by means of screw jacks until the SRP-HRL became 11.50 in. The legs were again retracted as far to the rear as possible.

When this posture had been held about thirty seconds and was reasonably relaxed and steady, the photographs were taken.

During the following period of camera loading, the subject moved his legs forward and the floor was again raised, to a SRP-HRL of 9.00 in. The legs were again retracted maximally and as soon as a relatively steady relaxed state was achieved, the photographic record was made.

This identical process was repeated for the SRP-HRL of 6.50 after which the subject was dismissed.

It was not considered necessary to re-examine the preliminary and ejection attitudes at each SRP-HRL level since preliminary studies had failed to show a significant change in any dimensions save those of the feet and leg clearances and lateral stick displacement as a function of the SRP-HRL settings used here.

It will be recalled that information descriptive of two phenomena was sought in this study: (1) the loci in space of certain anatomical features along with their associated clearance dimensions; and (2) the effect of limited changes in body configuration upon these. It is obvious that there exists between the two sets of parameters some sort of interdependence since dimensions are partially a function of body configuration and body configuration is partially a function of body dimensional parameters *per se*.

Together these two body attributes dictate dimensional clearances for any packaging and/or body support hardware for the human operator.

The resultant data on clearances and spatial loci measured on twenty-four subjects in each of the three principal body postures as well as

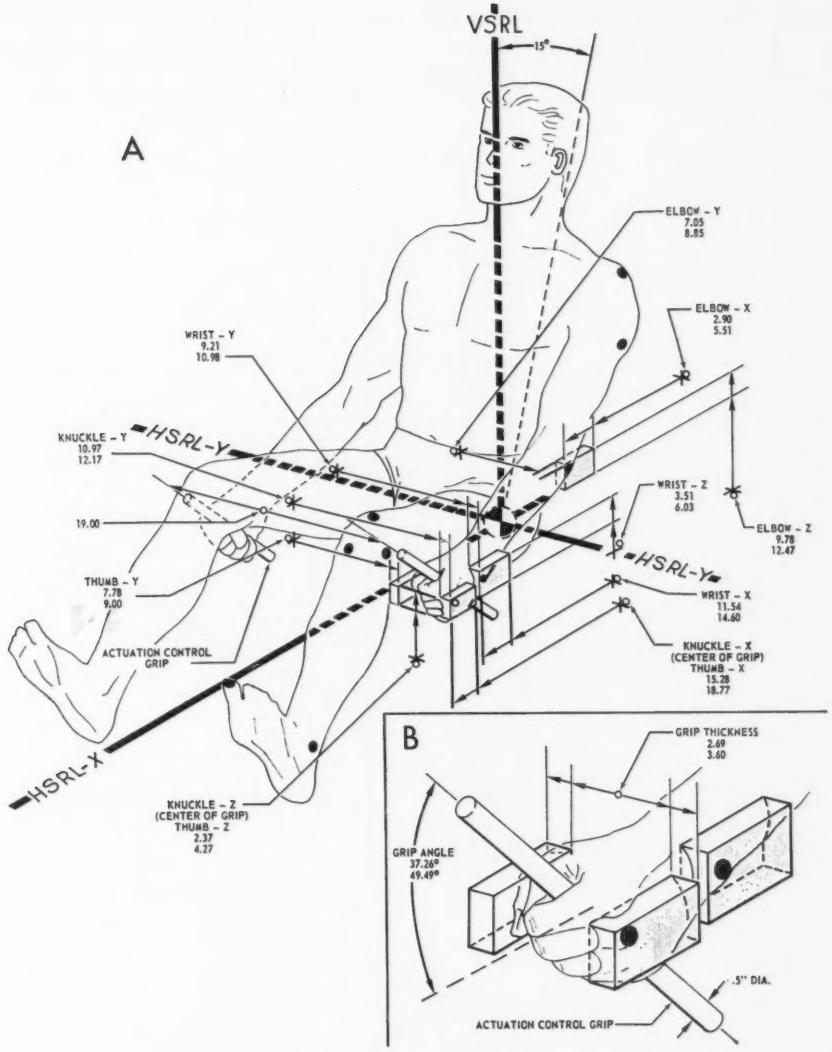


Fig. 6. Preliminary attitude.

the four leg geometry variants of the ejection and encapsulated flight attitudes, respectively, may most easily be presented in the graphic form of a series of pictorial summaries as shown in Figs. 6 to 11.

The numerical values shown on the figures represent the extimated lower and upper limits of the central 90 per cent of variability about the respective means of the dimensions shown. The lower limit (top number in each pair) is the 5th percentile, which is the mean minus estimated $\sigma \times 1.645$, and the upper limit (bottom number in each pair) is the 95th percentile or the mean plus estimated $\sigma \times 1.645$.

The use of these limits reflects the Air Force requirement to accommodate a *range* of human dimensional variability rather than an average value:

cannot satisfactorily provide the dimensions which will be suitable for the greatest number of USAF personnel."[3]

"As a general working principle in the design

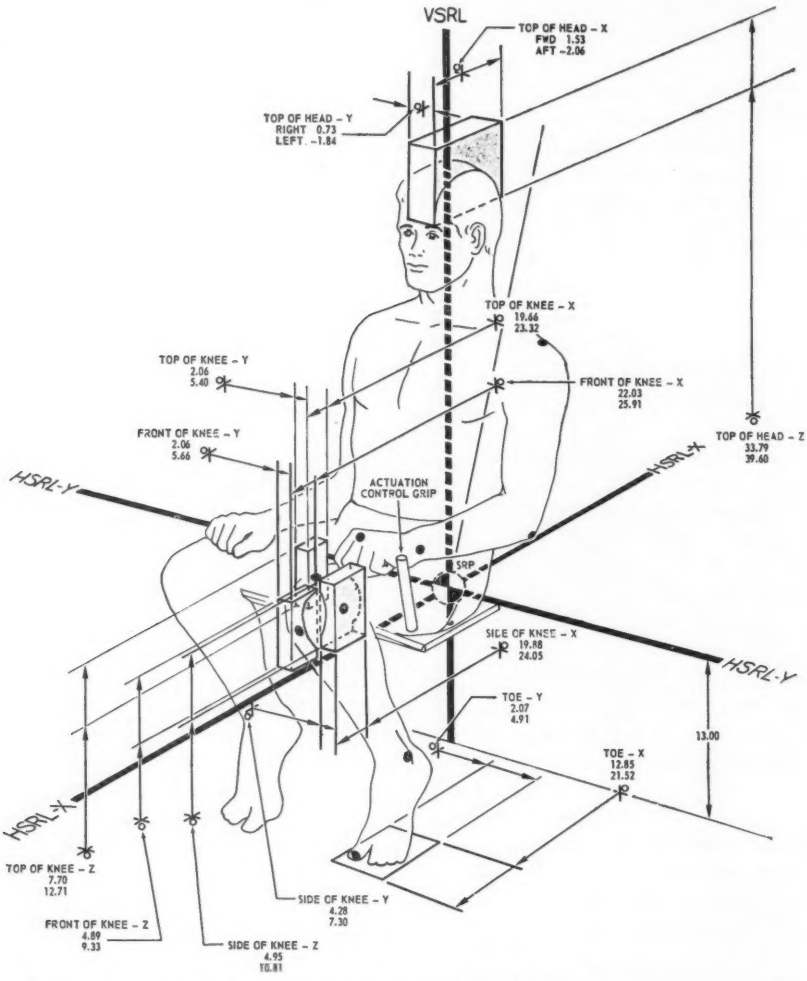


Fig. 7. Encapsulated flight (crew): SRP-HRL 13.00 inches. (5th and 95th percentile values.)

"The designers' immediate problem is that for all practical purposes, there is no such thing as the 'average man', reference WADC TM WCRD 53-7 and WADC TR 54-221. The practice of designing for the 'average' figure

of cockpits and personal equipment for USAF flying personnel, the range of accommodation is set as at least 90 per cent of the flying population, and more whenever possible. This percentage is chosen because of the great

expense of including the first few and last few per cent of the population. The 90 per cent range of accommodation previously mentioned uses the 5th percentile as its lower limit and the 95th percentile as its upper limit. The top and bottom five per cent groups are not ignored but cannot be accommodated as well because of the difficulty of providing full adjustability for them in aircraft."[2]

The question naturally arises as to which of the two values shown is to be used since an item of hardware, if it is not adjustable in size (for example, the shell of a capsule), must necessarily have but one fixed value for a given dimension. This question can only be answered after consideration of the function of the hardware. The shell of a capsule must utilize maximum dimensions (95th percentile) in certain areas in order to enclose the larger man.

The placement of the handgrips within this shell, on the other hand, must be based upon minimum or 5th percentile arm reach

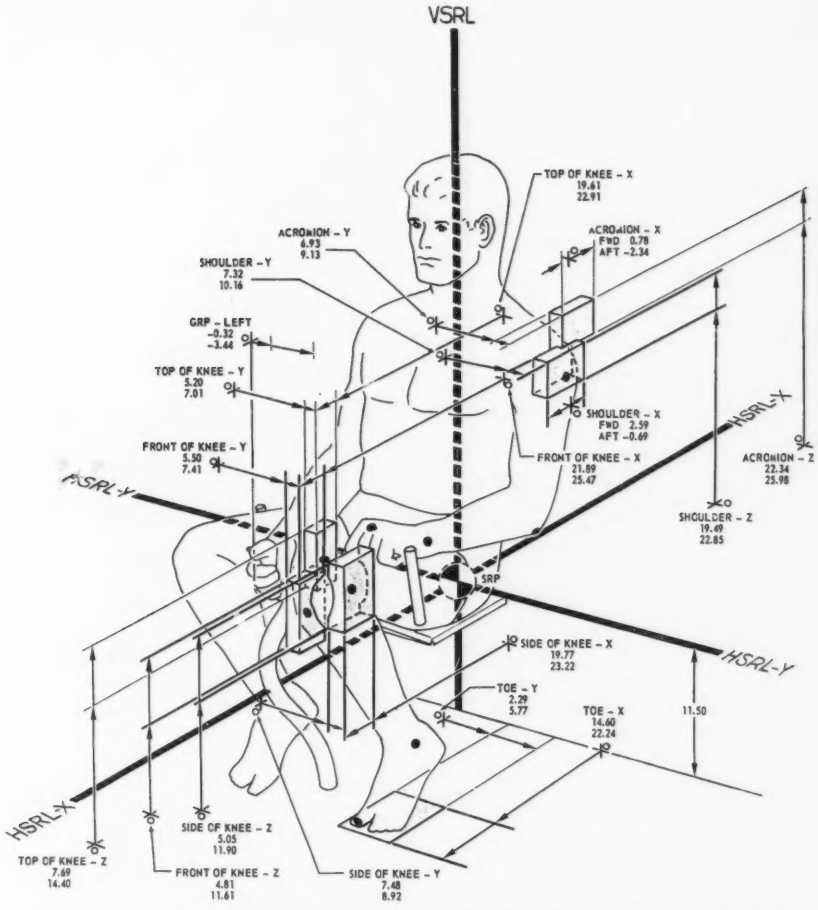


Fig. 8. Encapsulated flight (pilot): SRP-HRL 11.50 inches. (5th and 95th percentile values.)

dimensions to enable the smaller men to grasp them.

In order to illustrate the 5th and 95th percentile variability of the dimensional co-ordinates of the points measured, three-dimensional boxes have been used in the data presentation figures. In this manner is illustrated a spatial area within which 90 per cent of the observations of a given point occurred. If the range of variability necessary to be accommodated had been larger, say the 1st to the 99th percentile levels, all three of the dimensions of each area of each point locus would be appropriately enlarged.

As an illustrative example of the presentation technique, we may refer in Fig. 8 to the large block containing 90 per cent of the observed acromion points. It will be seen that the 90 per cent fore-aft or acromion$_x$ range of variability includes positions 2.34 in. aft of VSRL through positions lying 0.78 in. forward of VSRL.

These values correspond to the distance from SRP to the most forward (5th percentil value) boundary of the acromion area and to the

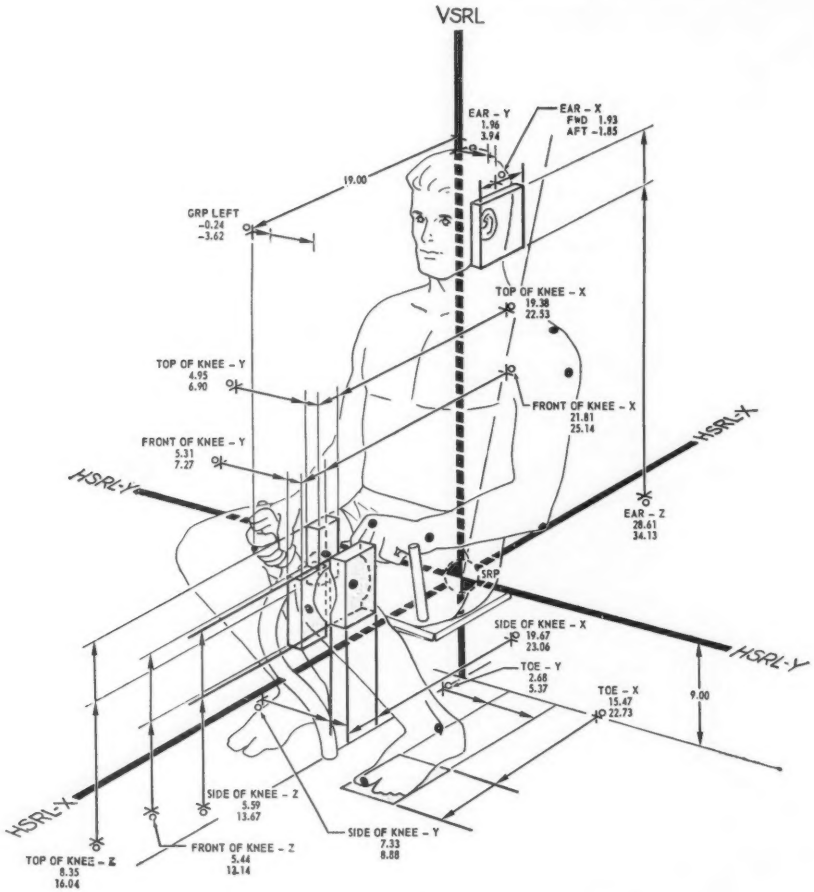


Fig. 9. Encapsulated flight (pilot): SRP-HRL 9.00 inches. (5th and 95th percentile values.)

distance from the SRP to the furthest (95th percentile value) aft boundary of the area.

Likewise, the 90 per cent lateral or acromion, range of variability of acromion includes positions 6.93 in. lateral to 9.13 in. lateral of the SRP. These values, in turn, correspond to the distance from SRP to the closest (5th percentile) lateral boundary of the acromion area and to the distance from SRP to the farthest (95th percentile) lateral boundary of the acromion area.

Finally, the 90 per cent vertical or acromion range of variability includes positions 22.34 in. above SRP to 25.98 in. above SRP. These dimensions represent the distance from SRP to the lowest (5th percentile) boundary of the acromion area and to the highest (95th percentile) boundary respectively of his same area.

In the data presentation figures (6–11) the origin of the various dimensions from the respective reference planes is indicated by a small cross and circle. The cross will be assumed to lie in the particular reference plane from which the dimension is called out.

A basic requirement of this study was that the data collected be reasonable estimates of similar parameters on the Air Force flying

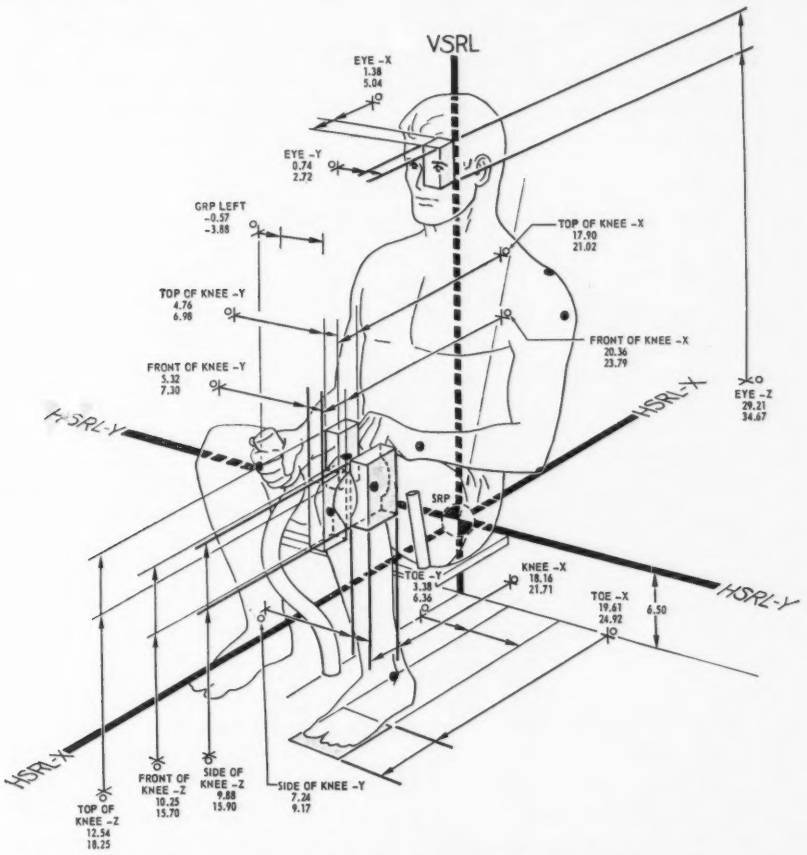


Fig. 10. Encapsulated flight (pilot): SRP-HRL 6.50 inches. (5th and 95th percentile values.)

population. Since it was impossible to actually determine these parameters on this entire population, a representative sample was required. To this end, twenty-four white male subjects were selected from among company (Convair-Fort Worth) engineering personnel for the photogrammetric study.

In order to estimate how well this sample represented dimensionally the Air Force flying population, thirty-one of the 132 measurements described in the 1950 survey[1] were taken on the twenty-four test subjects. As mentioned earlier, the only deviation from the technique used by the Air Force workers was the use of the Morant grid.[15]

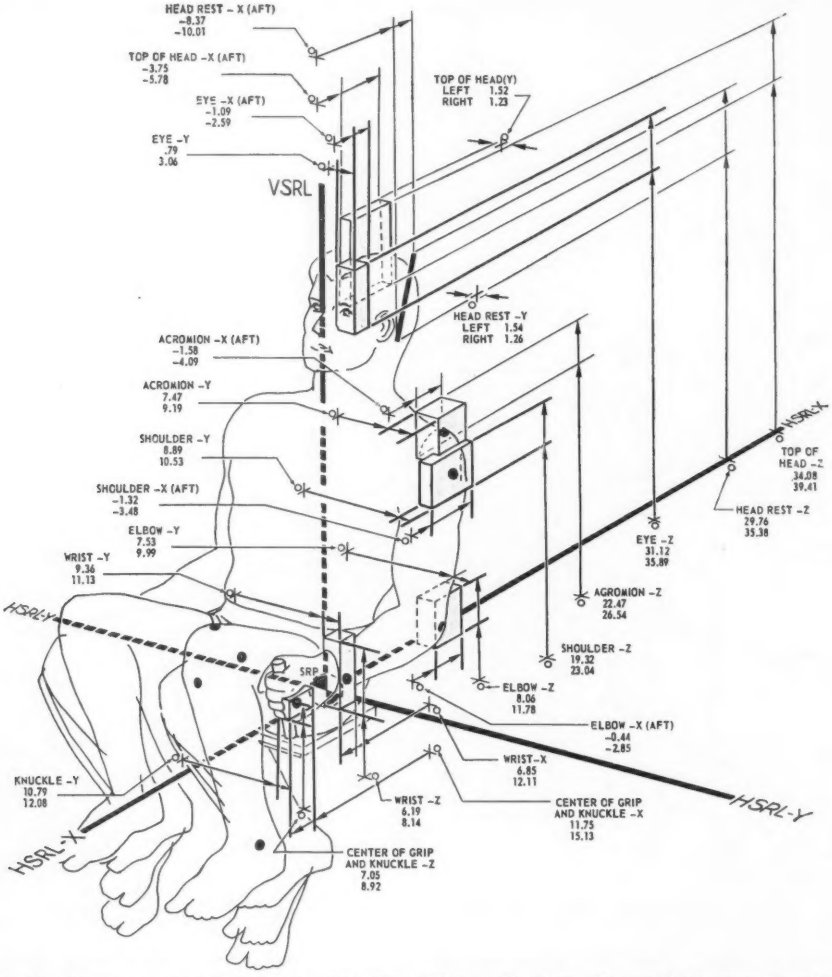


Fig. 11. Ejection (pilot and crew): SRP-HRL from 6.50 to 13.00 inches. (5th and 95th percentile values.)

In Table I, two statistical constants, the means and standard deviations (sigmas), describing ten of these thirty-one measurements

are arrayed next to the corresponding statistical constants taken from the report of the 1950 survey. The absolute differences are also listed.

Statistical tests were made to determine the significance of the observed differences between the two sets of means and between the two sets of sigmas. It should be noted that these tests, as described by Croxton and Cowden[6] are based upon a comparison between sample statistics (mean and sigma) and what are assumed to be known parameters of a population; e.g. the Air Force flying population. The hypothesis to be tested is that the sample could have been drawn from this population.

The observed differences between our test sample statistics and the 1950 survey parameters were evaluated in terms of the "odds" against getting differences of such magnitudes if the entire sample taken from the Air Force flying population in 1950 had been actually divided up at random into a large number of 24-subject samples and the differences between the constants derived from such hypothetical samples and the over-all population observed.

Rejection of our no-difference or "null" hypothesis at the 1 per cent level occurs if the observed difference between sample and population means and/or sigmas is of sufficient magnitude as to occur in less than 1 per cent of the above large number of hypothetical samples of the Air Force flying population.

In Table I, those observed differences having less than 1 per cent probability of occurring are marked by a *. Also shown, for means and sigmas respectively, are the magnitudes of difference required to go beyond the 1 per cent level of significance or probability.

It will be seen that in most cases the observed differences are well inside the 1 per cent boundaries. In such cases, the observed difference would be expected to occur due to chance factors much more often than 1 per cent of the time.

From these statistical tests, it is concluded, that while the means and sigmas of some of the Air Force "standard" dimensions recorded on our test sample, vary from their counterparts recorded on the Air Force flying population of 1950 by amounts sufficient to indicate possible non-randomness in sampling or even the sampling of a truly different population of such dimensions, others differ surprisingly little from this population considering the fairly high probability of getting even larger

TABLE I

Results of statistical tests of estimated constants: mean and sigma.

	Means				*Sigmas*			
Dimension	*Study Sample*	*1950 survey*	*Diff.*	*Diff. signif. beyond 1%*	*Study sample*	*1950 survey*	*Diff.*	*Diff. signif. beyond 1%*
1. Stature	69.96	69.11	0.85	1.28	3.03	2.44	0.59	2.25
2. Weight	170.16	163.66	6.50	10.97	30.17	20.86	9.31	19.21
3. Sitting height	36.40	35.94	0.46	0.68	1.62	1.29	0.33	1.19
4. Sitting eye height	31.74	31.47	0.27	0.67	1.45	1.27	0.18	1.17
5. Sitting acromion height	24.45	23.26	1.19*	0.60	1.25	1.14	0.11	1.05
6. Elbow rest height	9.65	9.12	0.53	0.55	1.01	1.04	0.03	0.96
7. Shoulder breadth	18.42	17.18	1.24*	0.48	1.14	0.91	0.23	0.83
8. Popliteus height	17.64	16.97	0.67*	0.40	0.99	0.77	0.22	0.71
9. Foot length	10.46	10.50	0.04	0.24	0.61	0.45	0.16	0.41
10. Chest circumference	37.78	38.80	1.02	1.29	3.64	2.45	1.19	2.26

* Significant beyond 1 per cent.

differences with random samples of twenty-four individuals each.

It may be reasonably concluded then that the "standard" dimensions of the study sample differed, by and large, no more from the Air Force parameters than might be expected due to chance factors and are, therefore, as representative as conditions will allow. It is difficult, however, to state the degree of confidence with which the same conclusions may be drawn about the new dimensions, without further analysis.

REFERENCES

1. ARDC, Cockpit Dimensions Stick Controlled Airplane (Excluding Helicopters), AD 1 Sheet in *Handbook of Instructions for Aircraft Designers*, tenth Edition, ARDCM 80–1, 15 June 1954 Supplement.
2. ARDC, *Handbook of Instructions for Aircraft Designers*, **1,** Part C, Chap. 2, Section 1.21., ARDCM 80–1, 1 April 1959 Supplement.
3. ARDC, *Handbook of Instructions for Ground Equipment Designers* (HIGED), ARDCM 80–5, Part A, Chap. 2, Section 7, Aug. 1957.
4. Chaffee, J. W., *Anthropometric Considerations for Escape Capsule Design*. Convair, Fort Worth, Texas, IFR 302, January 1960.
5. Churchill, E. and Bernhari, K., *WAF Trainee Body Dimensions: A Correlation Matrix*. WADC TR 57-197, WADC, W-PAFB, Ohio, 1957.
6. Croxton, F. E. and Cowden, D. J., Applied General Statistics, 2nd Edition, Prentice-Hall, Englewood Cliffs, N.J., 1956.
7. Davenport, C. B., Coordinates in Anthropometry. *Anthrop. Anz.* **3,** 1926.
8. Dempster, W. D., *Spatial Requirements of the Seated Operator*. WADC TR 55–159, WADC, W-PAFB, Ohio, 1955.
9. Gavan, J. A., *et al.*, Photography: An Anthropometric Tool. *Amer. J. Phys. Anthrop.* **10,** (3), 1952.
10. Hertzberg, H. T. E., *et al.*, *Anthropometry of Flying Personnel—1950*. WADC TR 52–321, WADC, W-PAFB, Ohio, 1954.
11. Hertzberg, H. T. E., *et al.*, Stereophotogrammetry as an Anthropometric Tool. *Photogramm. Engng.* **23,** (5), Dec. 1957.
12. Hooton, E. A., *A Survey in Seating*. Department of Anthropology, Harvard University, 1945.
13. Mason Electric Corporation, Grip Assembly, Type B-9 Control Stick, Drawing 22000, Los Angeles, California, 1954.
14. Miskin, E. A., The Application of Photogrammetric Techniques to Medical Problems. *The Photogrammetric Record*, **2,** (8), 1956.
15. Morant, G., Anthropometric Problems in the Royal Air Force. *Brit. Med. Bull.* **5,** (1), 1947.
16. Newman, R. W. and White, R. M., *Reference Anthropometry of Army Men*. U.S. Qmr. Corps., Environmental Protection Section Report (180) Lawrence, Massachusetts, 1951.
17. Randall, F. E., *et al.*, *Human Body Size in Military Aircraft and Personal Equipment*, TSEAA-695-43, Tech. Report 5501, AMC, W-PAFB, Ohio, 1946.

An Experimental Investigation of the Interaction Between Problem Load and Level of Training

RICHARD A. BEHAN, CHARLES R. BUGHMAN, JAMES N. BUMPUS and SALLY GILBERT,
System Development Corporation, Santa Monica, California.[1]

An experiment was made to determine the relation between performance in a linear information processing system, level of training and the input load on the system. Level of training was determined by number of problem experiences. Load was defined as the number of task units to be processed. The task was to detect, display, transform, re-display and correlate data concerning the temporal position and direction of sequences displayed on a rectangular grid. Two noise dots were included with each display presentation.

Experimental variables were presented in a mixed factorial design. Experimental problems had a constant number of sequences to be processed after a short build-up period. The sequences in any one problem appeared at regular intervals and all had the same time duration. Four experimental problems presented 2, 4, 8 or 12 sequences to the subjects.

Analysis of total error indicated that performance is a linear function of load, and that an interaction between total error and system position resulted in no significant reduction in error as a function of training.

A number of experimental studies have dealt with the effects of load on the performance of information-processing tasks[1, 2, 3, 4, 5, 6, 7, 8]. In general, the results of these studies have shown that the greater the load under which the system must operate, the greater the number of errors which will occur. Most of these studies were made in situations where the experimental subjects—human components of the systems—worked in only a single session of the experiment, i.e. a straightforward factorial experimental design was utilized. Such a study leaves unanswered a question which has importance from two points of view: (1) How much will training reduce the number of errors which occur under different magnitudes of load, and (2) What is the relationship between learning and performance under different magnitudes of load? The present study directs itself to the relationship between experience and performance under different conditions of load. In the present study, experience is defined in terms of the number of problem situations in which subjects have participated. Load is defined in terms of the number of task units to be processed per unit of time.

[1] The first three authors are now with the Thompson-Ramo-Wooldridge Corporation, DSL, Denver, Colorado. Mrs. Gilbert is in Dallas, Texas.

METHOD

Apparatus

The system used was a simple non-branching information-processing system consisting of five positions:

1. A reader, who organized an input and told information to
2. A first display position, which prepared an information display for
3. A teller, who told transformed information to
4. A second display position, which prepared an information display for
5. A correlator, who attempted to relate the display at position four with information from an independent source.

Stimuli

The initial input to the system was provided by a film strip projector. The projected display was in the form of dots on a rectangular grid. Each cell of the rectangular grid was named by a pair of numbers, e.g. 6, 7; 8, 2; etc. Square matrices of 100 cells were aggregated into larger cells which were named by letters of the alphabet, e.g. A, B, etc. The position of an input dot was determined by specifying the letter name of the large square and the number name of the small square in which it appeared

e.g. B, 4, 7; C, 7, 3; etc. A different display was presented every twenty seconds. The displays were systematically modified so that certain input dots (information) formed linear sequences.

The problems which were used in the study fell into one of three classes: (1) One problem which was used for orientation training and also as a warm-up problem on each of the experimental days; (2) specially constructed experimental problems; and (3) a single pre-test and post-test problem.

The orientation and warm-up problem contained six sequences in 36 twenty-second frames, and required twelve minutes to run off. Each subject was exposed to two runs of this problem during the orientation to his task. This, along with the discussion and demonstration which was individually administered, allowed the subject to become acquainted with his task. The pre-test and post-test problem contained thirty-six sequences in 120 twenty-second frames, requiring forty minutes to run off. All sequences proceeded as straight lines in one of the eight major directions of the compass. However, the sequences were of different lengths; hence, the load varied from one part of the problem to another.

The experimental problems were constructed so that there were always a constant number of sequences to be processed after a short build-up. The sequences in these problems appeared at regular intervals and all lasted for the same time duration. The appearance positions and directions were determined by reference to a table of random numbers. Load was determined by the number of sequences to be processed per unit of time. All of these problems consisted of 120 twenty-second frames. All sequences proceeded in straight lines in one of the eight major directions of the compass. Table I outlines the essential features of the experimental problems.

The task of the reader was to recognize and report the occurrence and the successive positions of these sequences of information. The positions of the dots on successive exposures were recorded. When three successive positions were occupied in the direction of any one of the eight major points of the compass, a sequence was formed. The reader then told the position and the time of the discovery of this sequence to the first display position. Each display presentation contained two dots which did not appear in any sequence. These were included to force the reader to filter the input. The reader told the position of the sequence at the time it was recognized, the clock time that it was recognized, the direction of progression of the sequence, and a sequence number which he assigned. Successive positions of each sequence were told at one-minute intervals.

The task at the first display position was to receive information about the positions of each of the sequences told him by the reader. These positions were recorded on a display board which contained, in addition to the rectangular co-ordinate system described above, a polar co-ordinate system. Sequence numbers, direction and time were retained throughout the system.

The task of the teller was to read the initial and successive positions of each of the sequences which was displayed on the first display board in polar co-ordinates and to tell this polar co-ordinate position to the system member at the second display position, thus completing the

TABLE I

Problems Used in the Experiment

Problem	*Appearance interval*	*Number of sequences*	*Sequence duration*	*Load*
Warm-up	Variable	6	Variable	Low
Pre-, Post-	Variable	36	Variable	High
Experimental problems				
BB4	4 min.	10	8 min. (24 frames)	2
BB3	2 min.	20	8 min. (24 frames)	4
BB2	1 min.	40	8 min. (24 frames)	8
BB1	40 sec.	60	8 min. (24 frames)	12

transformation of the information in the input from rectangular to polar co-ordinates.

The task at the second display position was to receive information about the successive positions of each of the sequences told him by the teller, and thus to prepare a display for the correlator. These positions were recorded on the second display board which contained only polar co-ordinates.

The task of the correlator was to compare selected positions of the sequences as displayed on the second display board with pre-plotted positions of potential sequences as prepared by the correlator from information presented on individual cards. Each sequence which appeared on the second display board was to be placed in one of two categories—those which matched or those which did not match one of the pre-plotted sequences. A record of the correlator's decisions was kept on a prepared form.

Design

The experimental design required eleven days for completion. The first day was devoted to orientation and one training run using the thirty-six-frame problem. The second day was devoted to a second training run with the thirty-six-frame problem, and closed with the run of the pre-test problem. Days three through ten were experimental days. On each of these days the thirty-six-frame problem was run as a warm-up problem, followed by one of the four experimental problems. Day eleven was devoted to the run of the post-test problem, which was preceded by the warm-up problem. See Table II.

The experimental problems were presented to the groups in a manner which was designed to cancel out effects of the order of problem presentation. Each problem was run on two successive days. Each problem-day-pair combination appeared before and after each other problem-day-pair combination an equal number of times. Such an arrangement with four problems required twelve groups.

Two hypotheses were tested: Difficulty, as measured by total error, is (1) a linear function of load, and (2) a linear function of the number of problems experienced by the group.

Just prior to each pre-, post-, and experimental run, group members were assembled in a separate experimental room. A member of the experimental staff presented a summary of the number of sequences which were properly handled at each position in the system on the previous day's run. This summary served to provide the group members with knowledge

TABLE II

The Complete Sequence of Runs for each of Twelve Crews (S—Shakedown, Pr—Pre, Po—Post, with Numbers indicating Problem Load)

	Days Training										
Crew	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1	S	Pr	2	2′	8	8′	12	12′	4	4′	Po
2	S	Pr	4	4′	12	12′	8	8′	2	2′	Po
3	S	Pr	12	12′	2	2′	8	8′	4	4′	Po
4	S	Pr	8	8′	4	4′	12	12′	2	2′	Po
5	S	Pr	12	12′	8	8′	4	4′	2	2′	Po
6	S	Pr	2	2′	4	4′	8	8′	12	12′	Po
7	S	Pr	8	8′	12	12′	2	2′	4	4′	Po
8	S	Pr	4	4′	2	2′	12	12′	8	8′	Po
9	S	Pr	8	8′	2	2′	4	4′	12	12′	Po
10	S	Pr	12	12′	4	4′	2	2′	8	8′	Po
11	S	Pr	4	4′	8	8′	2	2′	12	12′	Po
12	S	Pr	2	2′	12	12′	4	4′	8	8′	Po

of the results of their performance. The discussion period which followed was not structured by the staff member. He was there only to insure that the rules of procedure adopted for the experimental runs were maintained. The group members attempted to solve problems which arose at different positions during the exercises. There was no time limit placed on the duration of these briefings. When the discussion became non-task oriented, it was terminated.

Subjects

The subjects were male students from Santa Monica City College, who were naive with respect to the analog of the system used. These individuals were paid $1.50 an hour for each hour worked.

RESULTS

Frequency of errors occurrence was tabulated by system position for each of the pre-, post-, and experimental runs. Table III presents the analysis of variance for the total number of errors made during the eight experimental runs. All of the values of F which are greater than unity were significant at least at the 1 per cent level of confidence, except the F for experience which is significant at the 5 per cent level of confidence. There are significant differences in the number of errors made as a function of: (1) position in the system; (2) experience with problems; (3) load; (4) the interaction of load and experience; (5) the interaction of position and experience; (6) the interaction of load and position; (7) and the third order interaction between position, experience and load. Figures 1 and 2 present plots of total errors against experience and load, respectively.

Analysis of the variability of total error indicated that the contribution to variance of system position, load, and experience were significant. When these contributions to

TABLE III

Analysis of Variance of Total Errors

Source	*d.f.*	*SS*	*M.S.*	*F.*	*P*
Total	479	967,667.5938			
Between	59	280,567.8398			
Position	4	216,902.0898	54,225.523	46.84	< 0.01
Error	55	63,665.7500	1,157.559		
Within	420	687,099.7540			
Run	1	229.6333	229.633	< 1.00	
Experience	3	3,144.9167	1,048.306	2.44	< 0.05
Load	3	151,709.7988	50,569.933	117.64	< 0.01
RxE	3	965.7167	321.906	< 1.00	
RxL	3	553.3672	184.456	< 1.00	
ExL	9	10,090.2852	1,121.143	2.61	< 0.01
RxExL	9	2,704.0157	300.446	< 1.00	
PxR	4	152.5352	38.134	< 1.00	
PxE	12	22,791.4590	1,899.288	4.42	< 0.01
PxL	12	323,066.5742	26,922.214	62.63	< 0.01
PxRxE	12	3,650.2834	304.190	< 1.00	
PxRxL	12	1,238.6367	103.220	< 1.00	
PxExL	36	43,192.6797	1,199.797	2.79	< 0.01
PxRxExL	36	9,955.4785	276.541	< 1.00	
Error	265	113,654.3737	429.884		

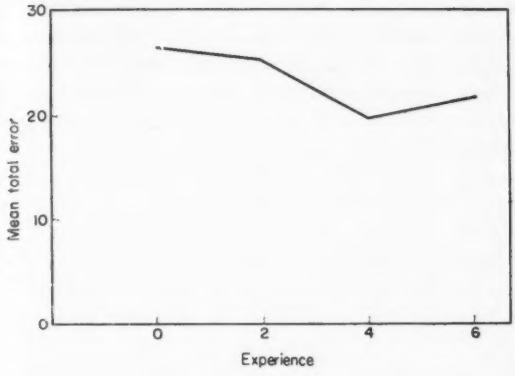


Fig. 1. Mean total error at each of four levels of experience.

TABLE IV

Analysis of Variance of the Number of Errors for the Pre-test, Post-test Runs by System Position

Source of variance	*d.f.*	*Mean square*	*F*	*P*
Total	119	60050.347		
Pre-, Post-	1	17040.833	23.02	< 0.01
Position	4	26462.321	35.74	< 0.01
PxP	4	15806.813	21.35	< 0.01
Error	110	740.380		

variance were tested against an error term which included the variability of the interactions involving these sources of variance, the contributions of load and position in the system were again significant beyond the 1 per cent level of confidence with F's of 32.30 and 13.77, respectively. The contribution of experience, however, was no longer significant at the 5 per cent level of confidence.

Table IV presents the result of the analysis

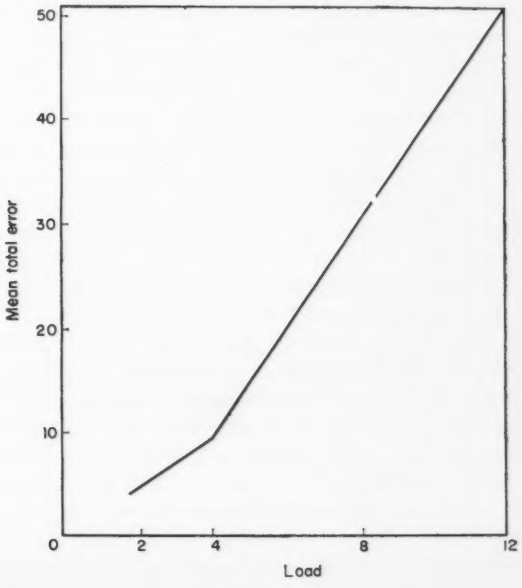


Fig. 2. Mean total error at each of four loads.

of variance of the number of errors made on the pre-test, and post-test runs. It is seen that the pre-test, post-test condition, system position and the interaction between pre-test, post-test and system position all contribute significantly to variance. The interaction between pre-test, post-test and position in the system is significant because of the relatively greater improvement which occurred in the first system position. There was relatively little change at other system positions. When the mean square for the pre-test, post-test condition and that for position was compared with an error term which contained the interaction of pre-test, post-test and position both were still significant beyond the 1 per cent level of confidence with F's of 13.43 and 20.85, respectively.

DISCUSSION

The relation between load and performance, as measured by total error, seems to approximate a linear effect very closely. (See Fig. 2.) This general statement was not, however, supported by a consideration of the different kinds of errors. It would seem that the relation between load and performance is a function of the measure of performance. Thus, while the present study supports the existence of the concept of load, it shows also that the usefulness of the concept will depend upon the adoption of some convention for the measurement of load. The results of the present study suggest that the convention should be total number of errors. Thus, hypothesis one is supported.

It might be remarked in passing that this selection of an appropriate measure for load in no way casts doubt on the validity of the concept. All techniques of measurement are conventions. Some techniques of measurement are so familiar to us that they seem to have a certain *a priori* characteristic. The reason that they are familiar and that they have survived for so long is that the conventions on which they are based were chosen to make the techniques maximally useful, and that these techniques for measurement are assumed in modern scientific theory. In other words, these measuring techniques are useful. The present finding of a linear relationship between load and performance when performance is measured by total error is most fortunate for it means that we are in fact dealing with a scalar magnitude and that we will be able to use the members of the number system as names of the magnitudes of load possessed by problems. The present finding is, however, only the first of many results of experiments which will have to be done if we are, in fact, to be able to measure load in general. The present experiment was conducted in a laboratory situation. Before load can be measured widely, experiments will have to be conducted to establish the relationship between our laboratory results and the results which might be obtained in the wider context. Secondly, our result was obtained using an analogy to the surveillance system of a manual Air Defense Direction Center. Any attempt to generalize these results to other functions of the same total system, or to the functions of other systems, will require further experimental support.

The overall effect of experience as demonstrated in the experiment was very small. (See Fig. 1.) The reasons for this are to be found by consulting the interaction between experience and position in the system. The only position at which improvement in performance occurred was the first, i.e. the input position. This is a function of the experiment itself and indicates that either the practice period was not long enough to allow the effects of experience to make themselves noticeable, or that the system used was not easily modified by training. The latter assumption is denied by results which were obtained during shakedown runs when the system was initially set up. During this period of shakedown runs, a group of girls from the Research and Analysis Group of the System Development Corporation served as subjects, and achieved a very respectable performance on a forty-track problem which resembled very closely the third level of load

as used in this study. The conclusion is that we did not allow enough trials for the effects of practice to become apparent, and that the experiment did not constitute an adequate test of hypothesis two.

SUMMARY

An experiment was made to determine the function which describes the relation between performance and problem load. Load was defined by the number of task units to be processed per unit of time. Performance was measured by counting the total number of errors which occurred in the system. Analysis of the total errors indicated that performance is a linear function of the magnitude of load.

REFERENCES

1. Christie, L. S., Luce, R. D. and Mace, J., Communication and Learning in Task Oriented Groups. *Technical Report* 231, Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Mass., May 1952.

2. Lindsley, D. B., *et al.*, A Study of Performance Under Speed Stress. Publ. Ed., No. 18381, 1945.

3. Lanzetta, J. T., Group Behavior Under Stress. *Hum. Relat.* **8,** 29, 1955.

4. Lanzetta, J. T. and Roby, T. B., Work Group Structure, Communication, and Group Performance. *Sociometry*, **19,** 105, 1956.

5. Lanzetta, J. T. and Roby, T. B., Effects of Work-Group Structure and Certain Task Variables on Group Performance. *J. Abn. Soc. Psych.* **53,** 307, 1956.

6. Lanzetta, J. T. and Roby, T. B., Group Learning and Communications as a Function of Task and Structure "Demands". *J. Abn. Soc. Psych.* **55,** 121, 1957.

7. McCurdy, H. G. and Lambert, W. E., The Efficiency of Small Human Groups in the Solution of Problems Requiring Genuine Cooperation. *J. Personality*, **20,** 478, 1952.

8. Shaw, M. E., Some Effects of Unequal Distribution of Information Upon Group Performance in Various Communication Outs. *J. Abn. Soc. Psych.* **49,** 547, 1954.

COED—A Device for the Experimental Study of Man-Machine Systems

R. H. JOHNSON, D. A. GORDON, B. BERGUM[1] and W. E. PATTERSON, Bendix Systems Division.

A description is given of an experimental facility for investigating man-machine system design problems. The facility is called the COED (Computer Operated Electronic Display). It combines a very large capacity cathode ray tube (Digitron) with a high-speed computer (IBM 704). The components, programming, and uses of the device are described.

Simulators, presenting in a simplified form the basic variables underlying a task, have long been used as training aids and as devices for human factor experimentation. Recently, at the Bendix Systems Division, a simulator facility has been developed for experimentation on very broad classes of man-machine system problems. Requirements for the device were stated as follows:

1. Extreme flexibility to simulate the display configurations and control characteristics of aircraft, helicopters, spaceships, and other man-machine vehicle systems.
2. Ability to present a wide variety of feedback dynamics linking operator inputs to the vehicles' responses.
3. Ability to respond to operator's commands without appreciable lag.
4. Ability to provide sufficient documentation of the display and control events to permit adequate experimental analyses.

The experimental facility constructed to these requirements is called the COED—Computer Operated Electronic Display (Fig. 1). It combines a large-capacity cathode ray tube (Digitron) and an IBM 704 high-speed computer. This device has been used in the development of an airborne missile system (EAGLE) and has contributed solutions to problems involving allocation of functions (man and machine), the development of display and control equipment, and the standardization of operational procedures.

1 Presently at the Hum RRO Unit, Fort Bliss, Texas.

COED CONSTRUCTION

A block diagram of COED is shown in Fig. 2. The COED system has the following components: A. Cathode ray tube and character generator which provides the display. B. Drum buffer and "Read Drum Continuous" package which functions as an electronic adapter between the IBM and cathode ray tube equipment. The Read Drum Continuous package permits direct access to the computer drum, so that a continuous view of the situation appears on the screen except during display updating. C. The control and indicator panels which the observer manipulates in response to the display. (In Fig. 1 these panels include tile push buttons, rotary switches, and a bowling ball cursor control); D. The input-output system, which translates control actions into "bit" computer language, and "Real Time" package adapter, which permits commands from the controls to affect the display without delay. These units are described in detail below.

A. Character Generator and CRT Display

The character generator and CRT display unit, which outputs information to the operator, includes two 19-in. cathode ray tubes mounted in an over/under arrangement. The character generator permits display of a repertoire of up to 64 characters or symbols. The current list of symbols used on the COED display is shown in Fig. 3. These characters are formed by a "stroke writing" technique which forms combinations of up to 16 line segments connecting points on a 3 × 5 dot matrix.

One 36-bit computer word, generated by the drum, commands the selection and positioning of a character on the Digitron display. Each word also commands character size, upper or lower scope location, blink rate, and type of modulation of a straight line from the previous, to the present character position. The modulation also involves the type of line (solid, dashed) which appears between the successive words.

The computer magnetic drum provides the output to the cathode ray display. The magnetic drum permits cathode ray tube display information to be continuously displayed until updated by the computer program. The drum rotates at 40 revolutions per second, and contains 2048 words.

B. Drum Buffer and "Read Continuous" Package

The drum buffer unit controls the read-out from the drum to the character generator and display unit. This unit permits the cathode ray tube to display up to 256 characters at a repetition rate at 40 c/s. It may also be set to show 512 characters at 20 c/s, and so on up to a limit of the entire 2048 characters in the drum field at 5 c/s.

Displays of considerable complexity can be

Fig. 1. COED simulator.

presented, where the requirement is merely that the display be updated when changes, sufficiently large enough to be discriminated by the cathode ray tube, occur. A display which can be generated well within the 40 c/s rate, is shown in Fig. 4. This figure, a hypothetical air traffic control display, requires only 83 words of the drum. Of these words, 51 are fixed and only 32 need be updated by the dynamic portion of the program.

allows six inputs per bit position, and two outputs per position.

D. Input–Output System and "Real Time" Package

A subsystem is required to receive operator commands and input data into the computer. This is accomplished by the COED Input-Output system, and the IBM "Real Time" package. This package might be termed an

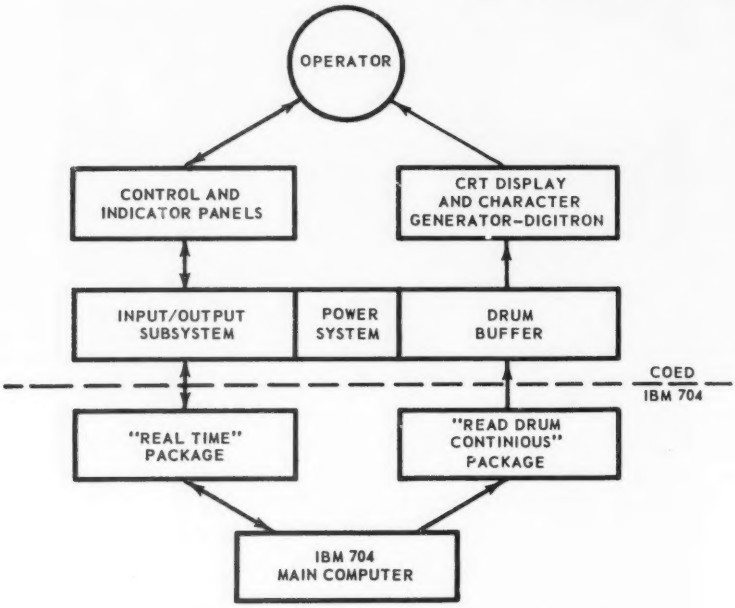


Fig. 2. Block diagram of COED system.

C. Control and Indicator Panels

Controls and indicators are mounted on patch panels, connected to plug boards in the wings (shelves of Digitron). The patch board permits experimental variation of the controls. A major change may be accomplished in a short period of time, since a great deal of the work may be done "off line". A switch panel can be fabricated, all hardware mounted, leads attached, and the proper timing signals inserted in the system by merely patching in these components. The present patching system

"actual" time device, since it permits the operator's commands to have priority over other inputs in affecting the display.

The Input-Output system accepts commands from a variety of input devices, varying in format for each COED control configuration, and provides an auxiliary output path for indications not desired on a cathode ray display.

In current applications, the Input-Output system is used to generate words on either program command, or following an interrupt signal generated by the Input-Output system

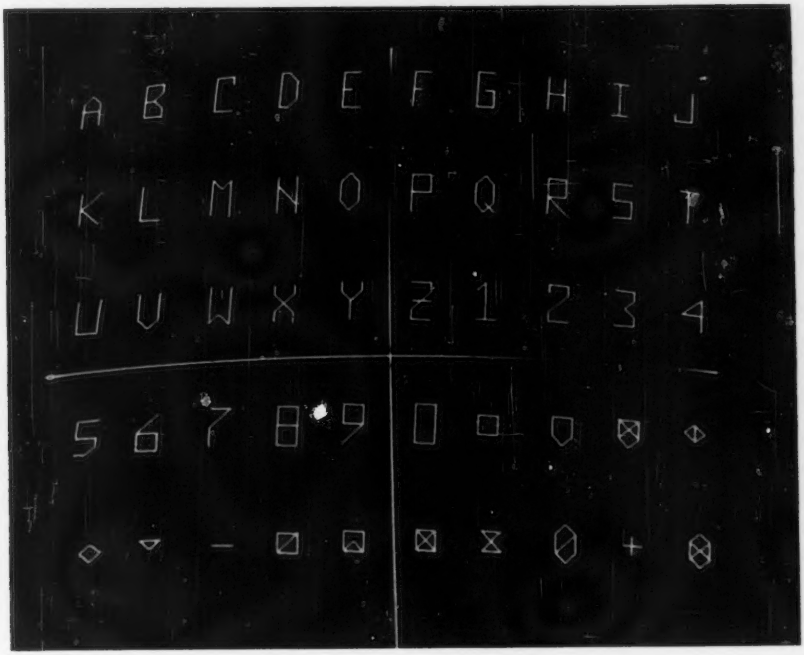


Fig. 3. Symbols used in display.

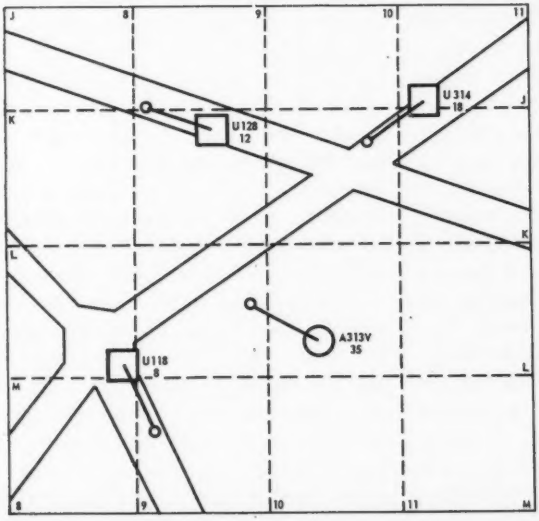


Fig. 4. Display requiring 83 computer words.

itself. The input words are normally coded by the input sense lines, and the computer then dissects these words to obtain the individual commands. The output words are similarly organized by the 704 program, and are routed to the correct indicators by the Input-Output logic.

APPLICATIONS

It has come to be realized, that the applications of a flexible display and command system and high-speed digital computer are far more numerous than first expected. Some of the applications of COED, in use or planned for use, are as follows.

A. Allocation of Function Problems

In designing a man-machine system, specific functions may be done either manually or automatically. An aircraft may be piloted or an autopilot used; an enemy target may be tracked manually or automatically. Many prescriptions have been given for allocation of system functions[2, 3] involving capabilities of man and machines, but these are often not useful in practice. And in the case of advanced systems, experience may not be available for detailing a satisfactory division of labor.

The COED permits an experimental approach to allocation of function problems. A manual system may be compared to one fully, or partially automatic. The radar beam may be programmed to seek incoming targets, or it may be set manually; piloting, selection of targets, and other systems functions may be performed in the two modes. In comparing manual with automatic operation, measures of accuracy and system time delays may be used as indices of performance.[1] Still another approach to allocation problems involves comparison of a simulated system against "ideal" performance. The detailed readout of the COED has been used for this purpose. Certain aspects of ideal performance may be easily specified—in piloting a plane, the ideal course is usually a straight line. In other instances, the time requirements for a sequence of operations may be stated. Where manual (or mechanical) performance is not adequate, the possibility of changing allocation to the other mode may be considered.

B. Display and Control Problems

The COED permits an approach to many classes of display and control problems, particularly where use of the device under simulated conditions will aid assessment. A "shakedown" technique has been found useful in dealing with display and control problems. In the shakedown, engineers who have designed the equipment carry out the simulated functions and state their reactions in a comprehensive questionnaire. Responses are coded to reveal areas of difficulty. Where alternate methods of setting up a control or display system exist, the choice can be resolved by formal experimentation involving dependent variables and conventional statistical methods of data treatment.

C. Problems of Face Validity

The realism required for a simulation is related to the problem under consideration. In the investigation of problems of functional allocation, the standard control panels (wings) are generally adequate. COED can also be used for face-valid simulations but at the expense of sacrificing some flexibility and increasing problem set up time. To achieve face validity, the CRT unit is remotable from the character generator section and the input-output console remotable from the Input/Output System. An example of the type of system simulation where face validity would be required, is the integration of a digital navigation and fire control system into an analog flight simulator. This type of simulation, although more costly than the non-face valid type, would represent an efficient integration of the best features of both analog and digital computers and of the display system. In this case COED would accept input data from the analog simulator as well as the pilot-operator.

SUMMARY

A description is given of an experimental facility for investigating man-machine system design problems. The facility is called the COED (Computer Operated Electronic Display). It combines a very large capacity cathode ray tube (Digitron) with a high-speed computer (IBM 704). This device may be used to simulate very broad classes of man-machine systems. It provides programmed feedback, in which the display changes as it would if controls of the actual system were activated.

The components of the COED simulator, as well as its programming and uses, are briefly described. The device may be employed to solve problems of allocation of function (man-machine), and those involved in the design of controls, displays, and procedures of use.

REFERENCES

1. Sinaiko, H. W. and Cartwright, G. P.,Careful: a pilot study of the effects of heavy target load on human and automatic decision makers. *Report R-215*, Coordinated Science Laboratory, U. of Illinois, Urbana, Ill., Sept. 1959.

2. Van Cott, H. and Altman, J. W., Procedures for Including Human Engineering Factors in the Development of Weapon Systems. *WADC Technical Report* 56–488. AD–97305, Oct. 1956.

3. Williams, A. C. Jr., Adelson, Marvin and Ritchie, Malcolm L., A Program of Human Engineering Research on the Design of Aircraft Instrument Displays and Controls. *WADC Technical Report* 56–526, July 1956.

Operational Sequence Diagrams in System Design

MARTIN I. KURKE, Dunlap and Associates Inc., Washington 4, D.C.

Operational sequence diagrams pictorially display information-decision-action sequences within a man-machine system. In its various versions as a time-sequence process chart, a spatial flow chart, and as an adjunct to symbolic logic, the OSD can be used in establishing system requirements, allocating man-machine functions, determining sequence of operations, and in evaluating equipment layouts.

DESCRIPTION OF THE OSD

During the past few years the staff of Dunlap and Associates has developed and refined a tool for the analysis of man-machine systems. This tool, called the Operational Sequence Diagram (OSD), is derived from methods engineering techniques such as the various types of operational process charts described by Barnes,[1] Mundell[8], and Maynard and Stegmerten.[6] These techniques have been supplemented by human factors concepts such as task analysis,[7] in which those discriminations, decisions, and actions necessary and sufficient to operate a mechanism are enumerated.

The Operational Sequence Diagram pictorially displays information-decision-action sequences a system or its component subsystems undergo to complete a mission. In an introductory paper Brooks[2] defined OSD's as schematic or decision-action flow diagrams and describes their use in the dynamic description of man-machine systems. He describes it as a convenient tool whereby

> ". . . system operation can be codified simultaneously with hardware designs. The human engineer can lay out his panels realistically; the project engineer can obtain a better idea of the man-machine relationships for various degrees of automation and can therefore evaluate alternative system designs."

The OSD may be used to establish sequence-of-operations requirements between subsystem interfaces at various levels of system analysis. The interfaces may be between machines, between operators, or between machines and operators. Simple or complex systems may be analyzed and the degree of thoroughness of the analysis may be selected for maximum usefulness. A second form of OSD stresses its use as a pictorial adjunct to symbolic logic. In this form the basic sequential OSD is combined with a logical analysis technique[5] to depict the logical result of each of several decision-action sequences. The third use of the OSD combines the technique with various link analysis techniques as described by Channell and Tolcott[4] for evaluating panel layout and the design of workspaces.

The basic components of an OSD are various geometric figures coded to denote the elements of any operational sequence (Fig. 1). In its simplest form the OSD consists of "action" elements depicted by squares, connected by a "time line" uniting the actions. The communication of information in the broadest sense of the term is indicated by a triangle to represent transmission, i.e. talk, switch closure, etc., and a circle to represent reception, i.e. hearing words, visual displays, etc. A semicircle denotes the utilization of stored information such as previously obtained knowledge or training.

When actions and communications are present, the next level of complexity involves the use of the element (hexagon) representing the decision function. A decision may be defined operationally as the variable intervening between a matrix of stimuli and a matrix of possible responses. It is the *raison d'être* for the OSD. The operator (which may or may not be human) requires information which is somehow integrated (the decision is made) and elicits one

of alternative actions or transmissions. As noted in Fig. 1, the resultant actions, inactions, or inappropriate actions can be coded in the OSD. Two additional types of symbol are

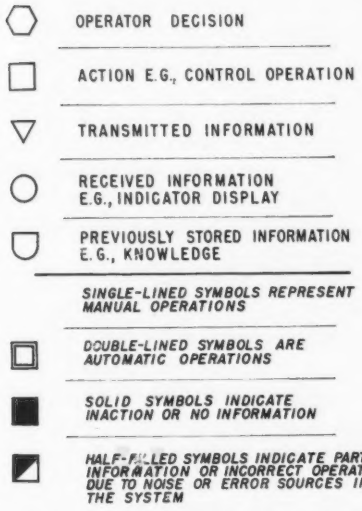

OPERATOR DECISION
ACTION E.G., CONTROL OPERATION
TRANSMITTED INFORMATION
RECEIVED INFORMATION
E.G., INDICATOR DISPLAY
PREVIOUSLY STORED INFORMATION
E.G., KNOWLEDGE
SINGLE-LINED SYMBOLS REPRESENT MANUAL OPERATIONS
DOUBLE-LINED SYMBOLS ARE AUTOMATIC OPERATIONS
SOLID SYMBOLS INDICATE INACTION OR NO INFORMATION
HALF-FILLED SYMBOLS INDICATE PARTIAL INFORMATION OR INCORRECT OPERATIONS DUE TO NOISE OR ERROR SOURCES IN THE SYSTEM

Fig. 1. Symbols used in operator sequence diagrams.

useful at this level of system complexity. One of these discriminates between manual and automatic functions. The other type is a logic device.

Manual elements such as those involving a man driving an auto, seeing a red light, deciding to stop, stepping on the brake, or blowing his horn at a pedestrian are represented by single-lined symbols. The brake light on his car is an example of automatically transmitted information. Automatic elements are depicted by double-lined squares, triangles, etc. For instance, the brake light of the car going on would be depicted by a double-lined symbol. The second type of symbol is logical in nature: it discriminates between logic conjunctions. Separate time lines entering or leaving an element represent an "or" condition. "And" is represented by the junction of two lines prior to entering, or subsequent to leaving the element.

USE OF THE OSD

Operator Sequence Diagrams may be used during various phases of system development from the initial determination of the sequence of gross operations to the detailed evaluation of panel layouts. To illustrate its use, let us examine how it might be used in the development of Campbell and Blank's proposal for a shipboard radar data computer to improve navigation safety by reducing the likelihood of ship collisions.[3] Such a collision avoidance system consists of a ship and an environment which contains objects to be avoided. To accomplish its mission the ship must meet the following requirements. It needs:

(a) Knowledge of potential collision-objects in the environment.

(b) Knowledge of own status relative to objects in the environment.

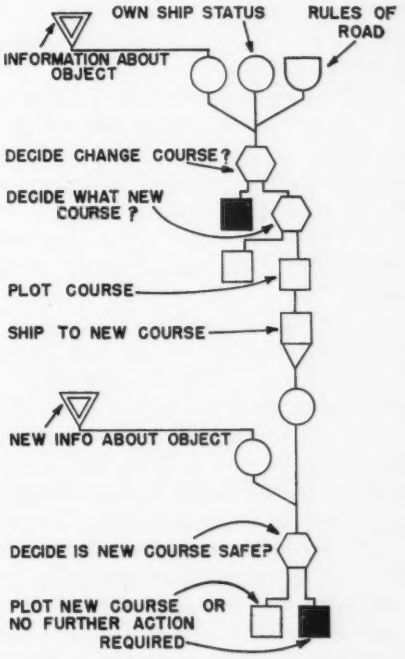

OWN SHIP STATUS
RULES OF ROAD
INFORMATION ABOUT OBJECT
DECIDE CHANGE COURSE?
DECIDE WHAT NEW COURSE ?
PLOT COURSE
SHIP TO NEW COURSE
NEW INFO ABOUT OBJECT
DECIDE IS NEW COURSE SAFE?
PLOT NEW COURSE OR NO FURTHER ACTION REQUIRED

Fig. 2. Basic Collision Avoidance System.

(c) Capability of determining CPA (closest point of approach), and from this deciding whether to evade the object.
(d) Ability to take actions necessary for evasion from a universe of acceptable evasive tactics.
(e) Feedback for evaluation of usefulness of evasive tactic.

Assembling these functions in a sequential order, the mission can be generalized as illustrated by the sequential OSD in Fig. 2, in which information concerning navigation hazards, such as another ship, is received by the collision avoidance system. This information, knowledge of own ship's heading, speed, etc., and of those actions permitted by the rules of the nautical road are all utilized in the decision to change own ship's course. A new course is chosen and plotted and the ship is put on the new course. The new relative bearing and speed of the other ship is received and the decision concerning the adequacy of the new situation is evaluated to determine the effectiveness of the maneuver.

The first step in developing the system is to compare the human factors elements in the existing system and in a proposed radar data computer-equipped collision avoidance system. An examination of the allocated man-machine functions can be made to yield comparative data on the human error potential of both systems.

The very question of the desirability of the hypothetical radar computer is one of allocation of man-machine functions. In effect the development of the radar computer would produce a new system and the relative effectiveness of the two systems should be compared. A very satisfactory method of determining the effectiveness of alternate systems is the comparison of OSD's representing them. More detailed diagrams than the one in Fig. 2 are needed. In addition to the ship and its environment, we are now concerned with the watch officer and the radar computer, and in the detailed interactions of these components. Fig. 3 illustrates this level of analysis of the conventional system and the proposed system incorporating a radar data computer.

The manual system (Fig. 3a) requires that a target be identified, and if a casual estimate indicates it might be a threat, its course relative to own ship is plotted to determine if the two ships are on a collision course. An evaluation of the plotted course is made and if the decision to maneuver to avoid the other ship is made, a new course is plotted. Then own ship is put on a new speed and heading. Subsequent to this change, the relative course of the other ship is re-plotted and the collision threat is re-evaluated.

The computer system (Fig. 3b) also requires that a target be recognized and identified. The target is then entered into the computer which continuously evaluates the relative threat. When the evaluation indicates a collision course an alarm notifies the watch officer who, on the basis of computer-fed displays, identifies which of several targets represents the collision threat. The watch officer simply estimates a change of course and/or speed and enters this change in the computer which evaluates it. If the threat is eliminated the alarm and threat indications are replaced by an "OK" display. The watch officer now simply orders the change in speed and heading he previously had entered into the computer. This entire sequence can be completed in a fraction of the time required for the manual sequence. In addition to graphically illustrating elapsed time, a time scale on the OSD shows the input and output rate load imposed upon the human operators in the system.

A comparison of the operations involved in both systems indicates that the types of error leading to collision potential are less easily designed out of the original system than the one with the computer, i.e. provided a correct entry is made, computation of CPA is less subject to error and more rapidly performed by electronic than human computers.

In general when the sequence reaches a major decision point the practice is to prepare

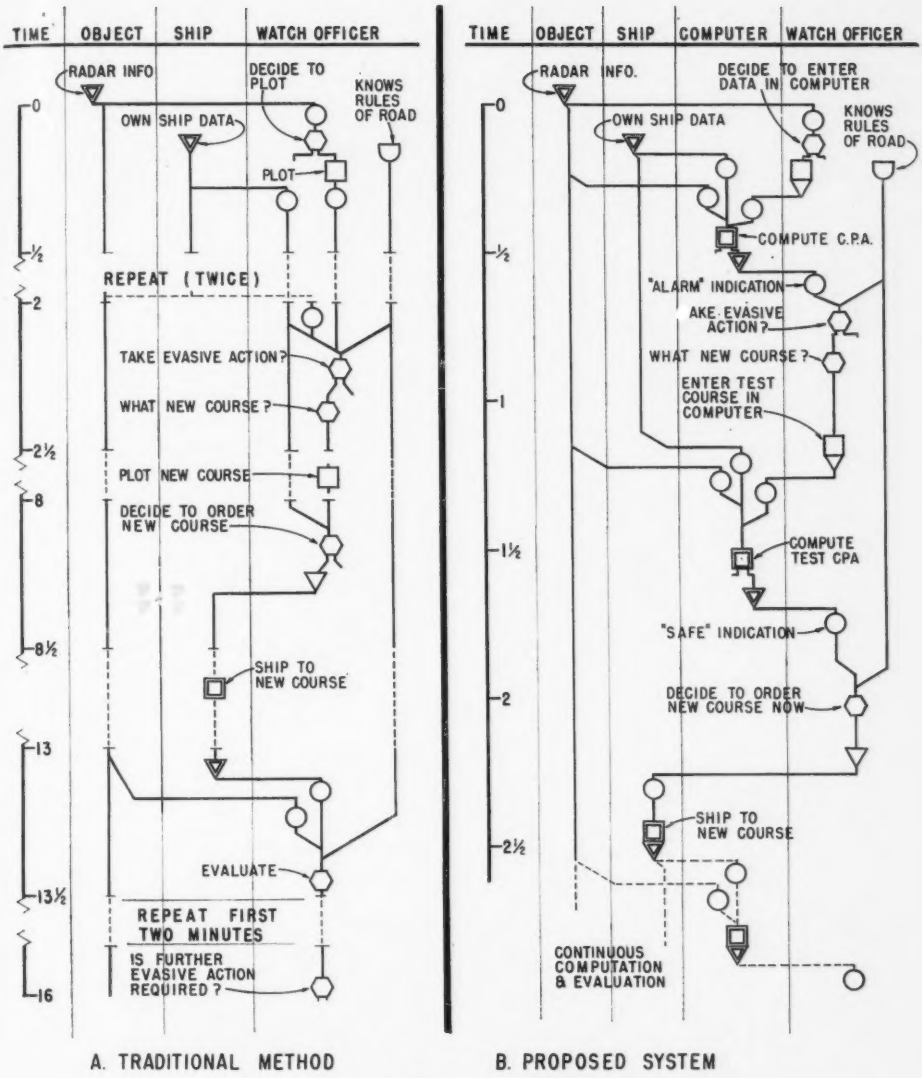


Fig. 3. Detail analysis of alternate collision avoidance systems.

separate OSD's for each alternate action. If a system requires numerous decisions this practice could result in more diagrams than can be handled conveniently. To overcome this disadvantage, computer programming techniques are incorporated into the OSD to show the effects of alternate actions on a single diagram. To illustrate the logical analysis form of the OSD, the watch officer in our hypothetical collision avoidance system makes decision *F*, the choice of a course (Fig. 4). Two alternatives, *G* and *H* are shown. If he enters course

G into the computer, data from the other ship via radar K and data concerning own ship L are integrated and computed. If conditions indicate that the entered alternate would lead to another collision course it would be so indicated by the computer's display R and a series of decisions (S, V) must be made by the watch officer to choose another course. If the

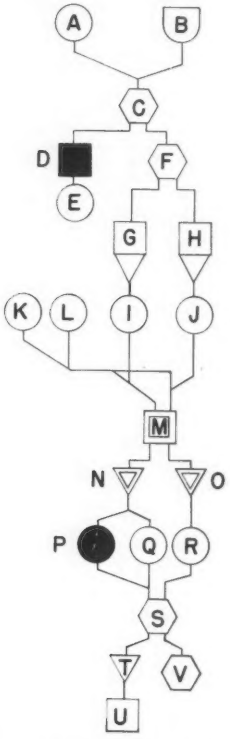


Fig. 4. Analysis of alternate actions.

alternative course H is chosen, this entry (with inputs K and L) will cause the indications P and Q that the entered course has a safe CPA. The watch officer uses this information in his decision S to order T, a change in the ship's course.

This use of the sequence diagram suggests its use with a system of symbolic logic. The results of the alternate decisions shown in Fig. 4 may be written as a series of logical sequences, thus:

(1) $A_i \cdot B_i \rightarrow C_i \rightarrow [(D_i \cdot F_o) \rightarrow E_i] + (D_o \cdot F_i)$

(2) $F_i \rightarrow [(G_i \rightarrow I_i) \cdot H_o] + [G_o \cdot (H_i \rightarrow J_i)] \rightarrow \bar{M}$

(3) $K_i \cdot L_i \cdot (I_i + J_i) \rightarrow M_i$

(4) $(H_i \rightarrow J_i) \cdot K_i \cdot L_i \rightarrow M_a \rightarrow (N_i \cdot O_o) \rightarrow (P_o \cdot Q_i) \rightarrow S_a \rightarrow T_i \rightarrow U_i$

(5) $(G_i \rightarrow I_i) \cdot K_i \cdot L_i \rightarrow M_b \rightarrow (N_o \cdot O_i) \rightarrow R_i \rightarrow S_b \rightarrow V_i$

where the capital letters stand for the elements in the OSD; small o used as a subscript represents the null state of the element; i is used to denote an only active state, and $a, b, c \ldots$, etc., are used as subscripts when an element has more than one active output; a dot indicates *and* logic and a plus indicates *or* logic. The bar over the capital letter means that the element indicated receives only part of its requisite inputs in this sequence. It indicates that other inputs to the element are necessary to change it from the null to the active state.

The ability to reduce graphic representation to the notation of symbolic logic enables the analyst to use the latter methodology in establishing meaningful operational procedures. This graphic-logical translation also provides a method for evaluating the relative effectiveness of various combinations of manned and automated system components. To illustrate this method of allocating man-machine functions, let us examine the hypothetical system shown in Fig. 5a in which act A results in a series of operations to elicit act K. Symbols B and C represent alternate environmental conditions which are necessary for, or inhibit this process respectively. B conditions occur 90 per cent of the time. Upon receipt of indications D and E, the operator decides (G) upon action H rather than J. H interacts with conditions B to produce act K. Should the decision G to activate H be made under environmental condition C, error L will result. Using the notation described earlier, the system under

normal operating conditions may be described as:

(6) $(A_i \rightarrow D_i)\cdot(B_b \rightarrow E_i) \rightarrow (G_a \rightarrow H_a \rightarrow \bar{K})+(G_b \rightarrow J_i)$

(7) $B_a \cdot H_a \rightarrow K_i$

Analysis of the diagram shows that the sources of "human error" are centered about the decision element G. The operator may respond incorrectly or fail to respond at all to various combinations of signals. In reference to decision G, the errors may be divided into input and output error, S and R, and described as:

(8) $S_i \rightarrow (D_o\cdot E_i \rightarrow G_i)+(D_o\cdot F_i \rightarrow G_i)+ +(D_i\cdot E_i \rightarrow G_o)+ +(D_i\cdot F_i \rightarrow G_a \rightarrow H_b \rightarrow \bar{L})$

and

(9) $R_i \rightarrow G_i \rightarrow (H_o\cdot J_o)+(H_o\cdot J_i)$

The human error can be reduced by two means. One method is by introduction of a "logic switch" to replace the human decision-making function (Fig. 5b). This will provide an automatized system where

(10) $D_i\cdot(E_i\cdot F_o) \rightarrow M_i \rightarrow H_i \rightarrow \bar{K}$

The other method would be the inclusion of additional components to permit the system to operate under condition C as well as B. The inclusion of component N would change the logic of the system to:

(11) $(A_i \rightarrow D_i)\cdot(B_b+C_a) \rightarrow [G_a \rightarrow (H_a \rightarrow \bar{K})\cdot(H_b \rightarrow \bar{N})]+ +(G_b \rightarrow J_i)$

(12) $C_b\cdot H_b \rightarrow N_i \rightarrow \bar{K}$

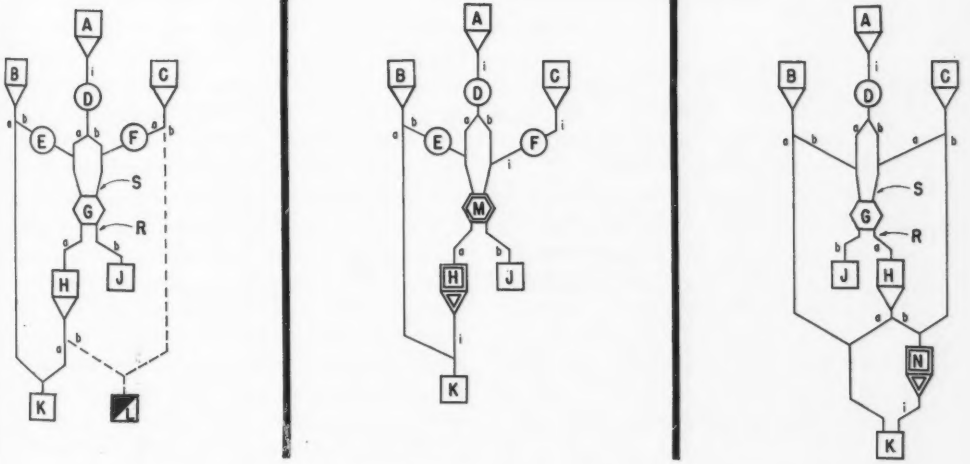
(13) $(B_a\cdot H_a)+(H_b \rightarrow N_i) \rightarrow K_i$

as shown in Fig. 5c. This arrangement would eliminate the need to discriminate between E and F. The nature of the input error source then becomes

(14) $S' \rightarrow [D_o\cdot(B_b+C_a)]+(D_i\cdot B_o\cdot C_o) \rightarrow G_i$

while R remains unchanged from equation (9).

The solution of the problem requires some quantitative information such as the reliability



A. ORIGINAL SYSTEM B. AUTOMATIZED VERSION C. MANUAL MODIFICATION

Fig. 5. System comparisons.

of components. For illustrative purposes let us assume the reliability of the original system which equals $(abcdef) = 0.788$, where

a = Reliability of Sequence $(A_i \rightarrow D_i)$ = 0.99
b = Ratio of B:B and C = 0.90
$c = 1 - P(S)$ = 0.95
$d = 1 - P(R)$ = 0.95
e = Reliability of Sequence $(B_a \rightarrow \bar{K})$ = 0.99
f = Reliability of Sequence $(G_i \rightarrow \bar{K})$ = 0.99

and $P(S)$ and $P(R)$ is the probability of human error at S and R respectively.

The first alternative of automatizing the system would replace human decision G with a logic switch M, thereby eliminating S and R. Assuming that the reliability of the logic switch is 0.99, this improvement would increase system reliability to 0.864.

The second alternative, of providing an additional component (with very poor reliability) to reduce the human error potential would yield a reliability of $(abdefh)+(adhjkmn) = 0.858+0.046 = 0.904$, where

$b = 1 - P(S')$ = 0.99
$j = C$:B and C = 0.10
k = Reliability of Sequence $(G_a \rightarrow H_b \rightarrow \bar{N})$ = 0.99
m = Reliability of Sequence $(C_b \rightarrow \bar{N})$ = 0.99
n = Reliability of Sequence $(H_b \cdot C_b \rightarrow N_i \rightarrow \bar{K})$ = 0.50

Collecting the analyzed data for various combinations of degree of automatization and component reliabilities as on Fig. 6 should yield data to establish the trade-off point when determining the allocation of man-machine functions in developing a system.

This particular analysis of system reliability leads to the conclusion that automatizing the existing system is considerably less effective than developing a new manual system with the capability of operating under a greater variety of environmental conditions. Some little additional improvement in system reliability can be obtained by automatizing the latter system. The data on Fig. 6, together with additional economic and engineering considerations (Schneider, 1960)[9] should be among the factors which determine how best to improve the system, i.e. whether to automate, to increase the scope of the system, to improve reliability of sequence $(C_b \rightarrow \bar{N})$, or some combination of these approaches.

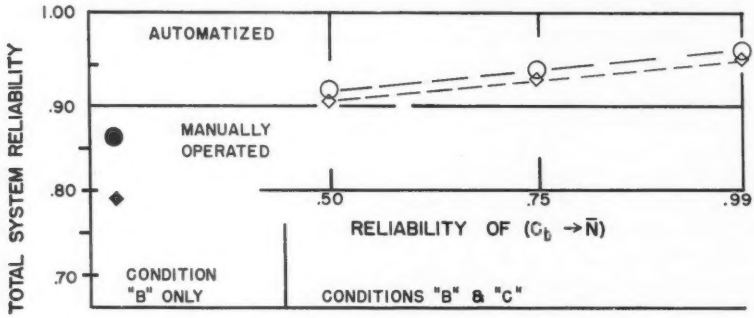


Fig. 6. System reliability.

In addition to its uses in system analysis, the Operational Sequence Diagram has its role in the design and evaluation of equipment. One method requires the development of very detailed OSD's showing on a time scale the inputs and outputs required by the operator for his information-decision-action sequences. These inputs and outputs may come from other equipment or from other personnel in the system. The time of input and output

information becomes one of the contributing factors in the design of equipment. The OSD used as a method for studying interpersonal communications required by the system can be used in the determination of the number and location of operating personnel.

Operational Sequence Diagrams may be laid out spatially as well as along a time line. The spatial layout method overlays the inputs and outputs to the operator on a sketch of the panel face. Fig. 7 illustrates this technique in the comparison of two panel designs submitted for a given sequence. Both panels consist of two toggles and six indicator lamps. The sequence of operations commences at the arrow and is shown by the line connecting three types of OSD elements: squares, standing for switch operations; circles for indicators on; and spots for indicators off. The spatial layout form of OSD provides a graphic description of the perceptual-motor load a particular layout imposes upon the decision-making function of the equipment operator.

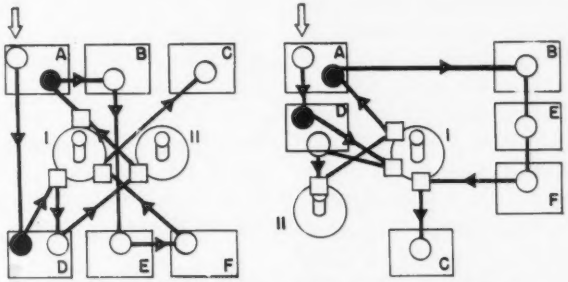


Fig. 7. Spatial OSD superimposed on two panel layouts.

SUMMARY

In summary, the Operational Sequence Diagram is a type of process chart modified for the peculiar needs of human factors work. Its primary use has been in determining man-machine interaction sequences, in analyzing communication requirements between groups of men and machines, and in coordinating information-decision-action sequences between interfacing subsystems. In its various forms the OSD can be used in several stages of system development. The initial and subsequent steps in establishing sequential system requirements may be aided by the OSD used as a sequence analysis tool. Both the sequential analysis and the logical analysis versions of the OSD may be used in the allocation of man-machine functions and in determining and evaluating the information-decision-action sequences and other procedures involving interaction between parts of a system. The sequence analysis and spatial analysis forms of OSD have been used to establish the input-output requirements of a panel or workspace and to evaluate such layouts.

REFERENCES

1. Barnes, R. M., *Motion and Time Study*. Wiley; New York, 1949.
2. Brooks, F. A., Jr. Operational sequence diagrams. *Inst. Radio Engrs. Trans. Human Factors in Electron.* **1**, 33, Mar., 1960.
3. Campbell, J. W. and Blank, H. E. Jr., Development of an experimental radar data computer. *Mar. News*, **46**, (6), 18, 1959.
4. Channell, R. C. and Tolcott, M. A., Arrangement of equipment. In: *Supplement to Human Factors in Undersea Warfare*. Washington, D.C., Natl. Acad. Sci.-Natl. Resch. Council, 1954.
5. Goode, H. H. and Machol, R. E., *System Engineering*. McGraw-Hill: New York, 1957.
6. Maynard, H. B. and Stegmerten, G. J., *Operation Analysis*. McGraw-Hill: New York, 1939.
7. Miller, R. B., A method for man-machine task analysis. *USAF, Wright Air Dev. Cent.*, Tech. Rep. No. 53-137, 1953.
8. Mundell, M. E., *Motion and Time Study: Principles and Practice*. Prentice-Hall: New York, 1955.
9. Schneider, R. H., Human factors in the establishment of system design requirements. *Presentation: Inst. Radio Engrs. WESCON: Conv. Rec.* 127, Aug., 1960.

The Authors

CHARLES W. SIMON (*Some Immediate Effects of Drowsiness and Sleep on Normal Human Performance*) recently became the human factors specialist at the Hughes Research Laboratories, Malibu, California, where he is investigating human pattern recognition and also the use of three-coordinate, integrated electronic displays. He entered the field of human factors in 1948 when he became the first Director of the Aviation Psychology Laboratory at Antioch College, supervising experiments on control-display relationships under contract with the Aero-Medical Laboratory at WADC. After receiving his Ph.D. in psychology from Ohio State University in 1952, he joined the Rand Corporation and participated in the early phases of the System Training Project. Later, he was the principle investigator relating the EEG at various levels of drowsiness and sleep to the ability to recall verbal material. In 1955, he joined the Human Factors Section of the Hughes Aircraft Company, Culver City, California, and conducted laboratory and flight experiments on the design of altimeters, magnifying optical sights, and radar displays. He also investigated ground radiation hazards of high-performance radar, long-range detections of Hughes radar systems, and factors related to aerial target identification. He was the first editor of the news *Bulletin* of the national Human Factors Society for two years, and is currently chairman of the Society's Public Information Committee.

ARTHUR I. SIEGEL and J. JAY WOLF (*A Technique for Evaluating Man-Machine System Designs*).

ARTHUR I. SIEGEL is currently the Director of Applied Psychological Services, Wayne, Pennsylvania. Dr. Siegel received the doctoral degree in psychology from New York University and for the past ten years has been engaged in visual, perceptual, human factors, performance evaluation, and aviation research for the Office of Naval Research, the Bureau of Aeronautics of the U.S. Navy, the Air Force, and for industrial clients. His experience includes the SAGE, ATLAS, 300A, CF105, F101B systems for the Air Force, and the P2V, F8U, S2F, ZPG systems for the Navy.

He has published over fifty professional articles and monographs and has presented sixteen papers before the American and Eastern Psychological Associations. Previously, he taught psychology at New York University and at Queens College, was a research associate at Fordham University, and a project director at the Institute for Research in Human Relations. Dr. Siegel is a member of the American Psychological Association, the Eastern Psychological Association, the Pennsylvania Psychological Association, the Human Factors Society, and the Aerospace Medical Association.

MR. J. JAY WOLF received the master's degree in 1951 in mathematics from the George Washington University, where he was later employed as Technical Director of digital computers on logistic research programs. Mr. Wolf's experience includes the solutions of mathematical problems related to the optimum allocation of functions, multiple criterion establishment, and maximization problems. Currently, he is involved as a research associate on an extension of the "model" described in order to enable computer simulation of multi-operator systems. Mr. Wolf is a member of the Society for Industrial and Applied Mathematics, the Research Society of America, and the Association for Computing Machinery.

ROBERT C. HAYGOOD and RICHARD K. OVERTON (*Operator Accuracy in Angle Measurement and Transfer*).

ROBERT C. HAYGOOD is a research engineer in the Human Factors Unit at Autonetics, a Division of North American Aviation,

Inc. His work there has been primarily concerned with experimental evaluation of alignment equipment and procedures for inertial guidance systems.

Mr. Haygood received his M.S. in psychology at the University of Utah in 1959. There he served as a USPHS Research Fellow and graduate research assistant, working on problems in human learning, prediction of scientific creativity, and programming for high-speed computers. He is a member of Sigma Xi, Psi Chi, and the Rocky Mountain Psychological Association.

RICHARD K. OVERTON is a senior research engineer in the Human Factors Unit at Autonetics, a Division of North American Aviation, Inc. His work there has influenced the design of optical equipment and devices associated with electronic computers.

For three years Dr. Overton was on the staff of Fort Hays Kansas State College, where he directed graduate work in psychology. He received his Ph.D. in psychology in 1956 from the University of Texas. His publications include several journal articles and industrial reports, and a book entitled *Thought and Action: A Physiological Approach.*

JOHN W. CHAFFEE (*Anthropometric Considerations for Escape Capsule Design*) has been chief of the Anthropometry Section of the Human Factors Laboratory at Convair, Fort Worth, since 1957. After receiving his B.A. in physical anthropology from the University of North Carolina in 1952, he worked as a project scientist in the Biokinematics Unit of the Anthropometry Section of the Aerospace Medical Laboratory of WADD from 1953 to 1956. His primary effort has been directed toward conducting studies of human body size, mobility, and muscular strength characteristics as related to work-station, escape systems, and ground-handling equipment design and operator task analysis. Mr. Chaffee's publications include *Anthropometric Considerations in Escape Capsule Design, A Study of Human Weight-Lifting Capabilities for Loading Ammunition into the F–86H Aircraft* (with I. L. Emanual), and *B–58 Optimum Brake Pedal Angle Study.*

RICHARD A. BEHAN, C. RUSSELL BUGHMAN, JAMES N. BUMPUS, and SALLY BONNEY GILBERT (*An Experimental Investigation of the Interaction Between Problem Load and Level of Training*). Richard A. Behan. See page 234, *Human Factors*, November, 1960, Vol. 2, No. 4.

C. RUSSELL BUGHMAN is a member of the Technical Staff at the Data Systems Laboratory of the Thompson Ramo Wooldridge Corporation at Denver, Colorado. He is currently involved with the evaluation of a large-scale information processing system. After receiving a B.A. in psychology from the University of Colorado, he joined the RAND Corporation. When the System Development Corporation was formed he joined the staff of the Research and Development Department where he participated in laboratory and field studies concerned with the organization and training of the Air Defense System.

JAMES N. BUMPUS is a member of the technical staff of the Data Systems Laboratory, Ramo-Wooldridge, Denver, Colorado. He is responsible for evaluation of some of the component and systems aspects of a large-scale data processing system under development for the Air Force. After receiving his B.A. in Sociology from the University of Colorado, Mr. Bumpus joined the RAND Corporation. When the System Development Corporation was formed, he became a member of SDC's Human Factors Research Staff. He directed a field study of certain organizational problems for the Air Defense Command. He also participated in a number of other laboratory and field research projects. He is a member of the American Sociological Association.

SALLY BONNEY GILBERT was born in Coos Bay, Oregon. She is a graduate of Scripps College, Claremont, California (B.A. 1952, Psychology) and of Stanford University (M.A. 1953, Guidance and Counseling). She joined the Research and Analysis Group of the System Development Division of the RAND Corporation in 1955 and transferred to the System Development Corporation when that company was formed in 1958. While at the System Development Corporation Mrs. Gilbert was a member of the Human Factors Research Staff. She was married in 1959 to Dr. Freeman Gilbert and they now live in Dallas, Texas.

PRINTED IN GREAT BRITAIN BY SPOTTISWOODE, BALLANTYNE & COMPANY LTD., LONDON AND COLCHESTER